Decoding Calcium Signals in an Early Diverging Land plant, *Marchantia polymorpha*

Connor Tansley

100089213

Doctor of Philosophy

University of East Anglia

December 2021

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there-from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

Calcium is a ubiquitous secondary messenger in plants that carries signals for a diverse range of stresses and developmental processes and coordinate responses to these stresses or developmental changes. Proteins decode these diverse signals through variable calcium binding domains. These decoder protein families have expanded extensively across the green lineage, resulting in functional redundancy between proteins which has made it hard to ascribe function to individual proteins. *Marchantia polymorpha* is a liverwort that diverged from the green lineage 450 million years ago and has not gone through any genome duplication or hybridisation events, making it a useful model for studying complex gene families.

This work shows that *M. polymorpha* has a simplified network of calcium decoder proteins including Calcineurin B-Like proteins (CBLs), CBL-Interacting Protein Kinases (CIPKs) and Calcium Dependent Protein Kinases (CDPKs). These calcium decoders fall into diverse subgroups which cover all of the subgroups of these protein families, likely meaning the decoders together cover many of the functions in angiosperms. Of these proteins, a subset show transcriptional changes in response to salt stress. Protein-protein interaction tests reveal that all the *M. polymorpha* CBLs and CIPKs are capable of interacting with each other. Genetic knock-out of a CIPK results in increased sensitivity to salt stress. A proposed CBL-CIPK pathway therefore likely forms part of a salt tolerance pathway in *M. polymorpha*.

Decreased functional redundancy in *M. polymorpha* make this system a good model plant system to investigate the unknown mechanistic details of the CBL-CIPK pathway and calcium decoding.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

List of Figures

List of Tables

List of abbreviations

AM: Amplitude modultation, FM: Frequency modulation, IP3R: Inositol triphosphate Receptors, VDCC: Voltage Dependent Calcium Channels, MSL: Mechanosensitive Ion Channels, TRP: Transient Receptor Potential Channels, P2X: Purinergic P2X channels, Cycs-Loop: Cysteine Loop Channels, ABA: Abscisic acid, CaM: Calmodulin, CML: Calmodulin-Like, CBL: Calcineurin B-Like, CIPK: CBL-Interacting Protein Kinase, CDPK: Calcium Dependent Protein Kinase, CCaMK: Calcium Calmodulin Dependent Protein Kinase, PPI: Phosphatase interaction motif, CaMK: Calmodulin Dependent Kinases, SOS: Salt Overly Sensitive,

Acknowledgements

Carrying out my PhD has been an amazing experience for me and has given me access to some of the cleverest and most caring mentors and friends. Firstly, I would like to thank Ben Miller who has been a mentor for me since my undergraduate degree and has calmly and consistently helped me through the experience with unending patience and enthusiasm. He constantly pushes me to carry out better work but also offers support when the experience has been hard. I cannot imagine having a better supervisor for the whole PhD experience.

Secondly, I would like to thank Taoyang Wu. Taoyang has constantly added insight and encouragement through the project and has always made me feel like I can do this, even when I was unsure. I have developed so much during every conversation we have had and he has really pushed my computational skills and mindset throughout.

I would also like to thank my funders, including the University of East Anglia and the UKRI BBSRC NRP DTP. The funders and the infrastructure they have built has provided me with constant training throughout the PhD and helped me to experience the full breadth of what a research career can give me. This research would not have been possible without the training and excellent mentors the teams have provided, as well as their financial support.

Lastly, I could not have done this research without my friends and labmates. Rocky Payet has been a particular source of fun and amusement while also being incredibly professional and willing to bounce wild scientific ideas around almost every day throughout my PhD. Mellieha Allen has constantly been an emotional

support since she joined the lab whether as a listening ear or as someone to have a laugh with when experiments were failing. James Houghton, Jack Stubbs and Lizzie Payne have been excellent students and seeing their development has been a wonderful encouragement. I have no doubt that they will go on to succeed at whatever they set their mind to. I am most grateful for the love and support of my partner Matthew Lewis for his extreme patience during the PhD and especially during the writing process. I am also very grateful for the fun times my friends Alan Houghton, Lauren Mills, Summer Rosonovski, Conor McGrath, Gracie Younie, Jacob Gretton, Tess Carr and Ruari Armstrong have given me over the years. This group have pulled me out of the work and constantly reenergised me throughout and it would not have been possible without them.

Preface

Calcium is used by a wide range of organisms as a secondary messenger for a range of responses to external conditions and developmental processes. Calcium is generally kept at lower concentrations in the cytosol than both intracellular stores and extracellular spaces. This is because calcium is capable of chelating phosphate which is integral to a number of processes in cells such as energy storage and use, DNA replication and signal transmission. A wide range of organisms have hence developed ion channels to release calcium into the cytosol and this calcium can encode messages to coordinate responses dependent on the dynamics of the calcium released. The calcium can vary in terms of amplitude (or concentration), duration (or length of calcium release) and frequency if the calcium release takes place more than once (the period between calcium releases) (Allan et al. 2022). Plants have demonstrated a wide array of stimuli can induce different dynamic calcium releases to a variety of abiotic stresses such as high external salt, low nitrogen, or drought (Knight et al. (1997), Liu et al. (2017)). Different dynamic signals are detected in response to different stresses which coordinate plant responses to the stress detected.

Calcium binding proteins are responsible for coordinating plant responses in response to different calcium signals. Many proteins have domains capable of binding calcium including the EF hand domain which directly binds calcium with different EF hands having different affinities for calcium which allows them to respond to different signals. Most proteins with EF hand domains have more than one domain capable of binding calcium which allows the protein to fine tune the calcium signal it will respond to (Christodoulou et al. 2004). Another factor which allows for further

signalling complexity is spatial segregation of signals, whereby different signals may take place in different tissue or organs or even in different intracellular spaces from different stores (Quan et al. 2007). Many calcium binding proteins are therefore expressed in specific tissues and have localisation domains that target them to different intracellular locations so that they are nearby to the ion channels that encode the signal that they respond to (Quan et al. 2007).

In angiosperms (flowering plants) many studies have been undertaken to understand the molecular and biochemical basis for decoding these calcium signals as well as ascribing functions to different calcium decoders. Thus far, several calcium binding affinities have been described for different calcium binding domain demonstration the specific nature of calcium decoders and crystal structure have demonstrated how the binding of calcium typically causes changes in conformation demonstrating the physical changes brought on by calcium signals (Sánchez-Barrena et al. 2013). Alongside this different calcium decoders in the same family of calcium decoders have been demonstrated to have different tissue or subcellular localisation (Quan et al. 2007). Also, different functions have been described even within single protein families including function in response to different ion stresses (Li et al. 2009). However due to genetic complexity in plant systems, with a wide range of hybridisations and whole genome duplications, many of the decoders of calcium signals have yet to be described (Hetherington and Brownlee 2004). Whole genome duplication has led to functional redundancy in decoder systems so that decoders have either the same function or overlap in function. Hence, single gene knockout does not lead to detectable phenotype due to compensation of functionally redundant decoders. A number of groups have attempted higher order genetic knockout informed by phylogenetics, but it can be difficult to propose which decoders in a particular family are functionally redundant and the proteins of interest may merely have overlapping function and have additional discrete functions which cannot be parsed using this method (Liu et al. 2017).

The work in this thesis aims to use an evolutionary approach to tackling the issue

with functional redundancy utilising a new model organism, *Marchantia polymorpha*. *Marchantia polymorpha* has recently had its genome published and demonstrates a lack of functional redundancy in a range of signalling pathways due to a lack of whole genome duplication in its evolutionary history (Bowman et al. 2017). Several methods have also been published to generate knockout mutants and store genetic lines and other key techniques to allow different signalling pathways to be studied (Sugano et al. (2014), Takahashi and Kodama (2020)). As a sister clade to land plants that diverged some 450 million years ago *M. polymorpha* may have a range of calcium signalling responses that are novel but it may also be able to shed light on the functions of a number of calcium decoders that have proven difficult to ascribe function to in angiosperms. Therefore this thesis will focus first on identifying a number of the calcium decoders genetically encoded in *M. polymorpha*, and then using expression profiling, protein-protein interactions and genetic knockout to ascribe function to a calcium decoder in response to salt stress in *M. polymorpha*. From this, evolutionary analysis will be used to compare this calcium decoder to those found in angiosperms such as *A. thaliana* to determine if the *M. polymorpha* system can be used to propose ancient functions to groups of decoders found in angiosperms.

Chapter 1

Introduction

1.1 Calcium as a Secondary Messenger

Release or influx of calcium at different subcellular and tissue localisations acts as an essential secondary messenger in animals and control processes such as contraction of the heart, formation of memory and cell morphology (Edel and Kudla 2015). In plants, calcium signalling is also important in plants for cell polarity, directional growth, fertilisation and responses to abiotic and biotic stresses (Boudsocq et al. 2010; Li et al. 2006; Boudsocq and Sheen 2013; Harper 2001; Knight and Knight 2001). In yeast, calcium signalling is involved in a range of processes including cold shock responses and iron toxicity (Peiter et al. 2005b).

It is thought that calcium signalling arose because calcium can chelate phosphate ions which are essential components of the DNA backbone, ATP, and intereference with phosphate donors such as ATP can inhibit kinase activity in cell signalling cascades. Mechanisms to extrude toxic calcium ions from the cell are found in a wide range of organisms and reduce calcium concentrations in cells beneath that of extracellular calcium concentrations. Due to these extrusion mechanisms, the concentration of intracellular cytosolic Ca^{2+} is maintained at 100-200 nM as opposed to the extracellular concentration which are commonly 20,000 fold higher (Clapham 2007; Wilson et al. 2015). Alongside this the cytosol further restricts calcium ion movement by reducing diffusion distance. In saline solution calcium ions can migrate

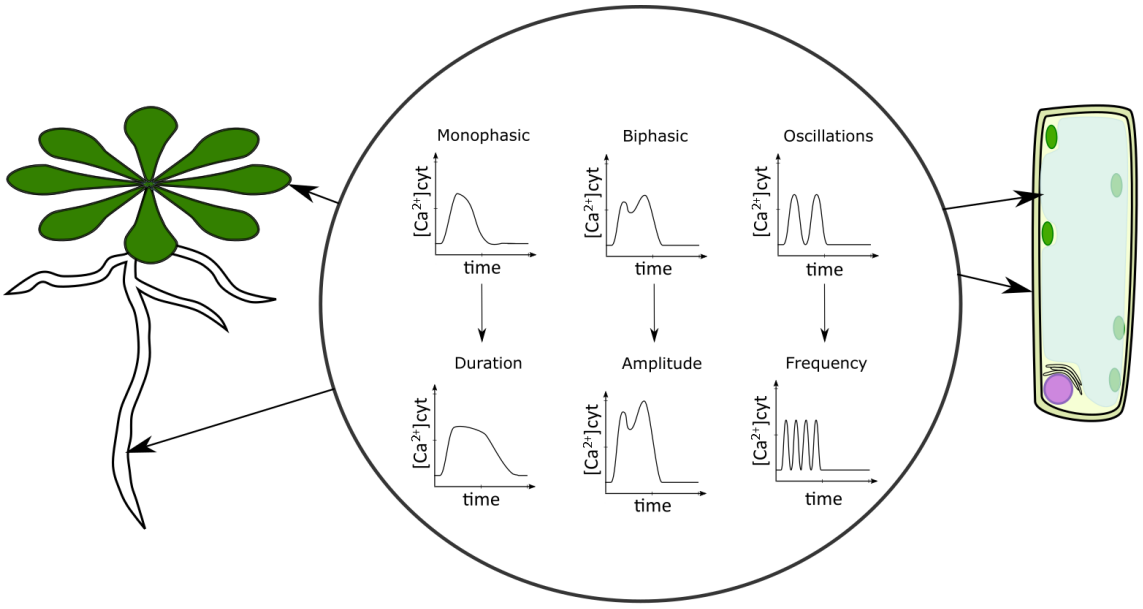


Figure 1.1: Model of a variety of calcium signatures. Top panel shows single step peaks (monophasic), two step peaks (biphasic), and peaks followed by a return to baseline followed by subsequent peaks (oscillations). Bottom panels show how signals can vary in length of the signal over time (duration), the amount of calcium released (amplitude: AM), and how many peaks occur within a set time period (frequency: FM). Arrows demonstrate the spatial segregation of calcium signals to different tissues including leaves (top left), or roots (bottom left). Different subcellular locations are also indicated such as the tonoplast (top right) or plasma membrane (bottom right)

40 µm/s whereas cytosolic calcium ion can only diffuse 100-500 nm before being extruded from the cell, and hence it only has an in-cell half life of 25 μs (Webb et al. 1996; Einstein 1905; Clapham 2007; Bowler and Fluhr 2000). This rapid extrusion against the concentration gradient contributes to the spatial restriction of calcium ions to specific subcellular compartments, or the outside of the cell, and limits the movement of the calcium ions when they enter the cell.

Calcium signalling commonly occurs in response to stress and occurs when stimuli open calcium ion channels allowing for the movement of calcium ions across a membrane. This results in an increase in local calcium ion concentration known as calcium influx. The calcium is then extruded, which is assisted by the reduced diffusion distance in the cell, as well as by active transport of the ions into intracellular stores, or outside of the cell, known as calcium efflux. The amount of calcium allowed into the cell determines the amplitude of the signal. The speed of extrusion

compared to the rate of influx, determines the duration of the signal. If the calcium ions are influxed and effluxed to basal concentration more than once this can be considered an oscillation. Oscillations have a frequency and amplitude which is analysed in terms of rate (i.e. the durations of each peak or frequency modulation (FM)) and total number of oscillations as well as peak size (or amplitude modulation (AM)). Single peaks before a return to basal calcium concentrations are considered monophasic. Biphasic peaks can also be seen whereby two peaks of calcium release and efflux overlap without a return to basal calcium concentration. The factors of amplitude, duration and frequency all affect the time and size of any peak in a calcium signal and are hence known as temporal restrictions (Fig.1.1). These temporal restrictions make different calcium signals known as calcium signatures and these signatures are specific to different stimuli and are determined by the calcium allowed in and the extrusion rate. Many of these temporal restrictions of the calcium signature are controlled by ion channels and therefore it can be said that ion channels encode the calcium signal (Hetherington and Brownlee 2004).

1.1.1 Calcium Signals are Encoded by Ion Channels

Calcium signals are encoded by release of calcium from cellular internal stores or from extracellular spaces across a membrane. Calcium ions (Ca^{2+}) cannot pass through the phospholipid bilayer of cell membranes due to their ionic charge. Therefore for signalling to take place calcium must be transported through the membrane. This is made possible through selective ion channels. One aspect of the complexity of calcium signalling is therefore defined by the spatial restriction of the signal, via the activation of specific calcium influx machineries, in specific locations in the organism and/or the cell. The machinery responsible for calcium influx varies from organism to organism. A range of different types of channels that allow passage of the calcium ions have been described in mammals: Inositol trisphosphate receptors (IP3Rs), voltage dependent calcium channels (VDCCs), mechanosensitive ion channels (MSLs), transient receptor potential channels (TRPs), purinergic P2X channels

(P2XRs) and cysteine loop channels (Cys-loop). However, it has been found that a number of these are not present in plants including IP3Rs, TRPs and four domain VDCCs (Verret et al. 2010; Wheeler and Brownlee 2008). The lack of some of these channels likely indicates that the machinery involved in encoding calcium signalling in plants has diverged from animals.

Voltage Dependent Calcium Channels

In animals VDCCs are specific channels that allow for the transport of calcium across the membrane. These channels are activated by changes in membrane potential local to the protein resulting in changes in conformation to allow for the movement of specific ions, in this case Ca^{2+}. VDCCs are specific to allow passage of calcium ions due to a conserved four aspartate residues in the pore forming p-loop domain (Yang et al. 1993; Iida et al. 2007). In yeast, one of these channels (Cch1pMid1p) is involved in responses to cold shock and iron toxicity (Peiter et al. 2005a). These channels are also involved in neurotransmission and regulation of gene transcription in humans and fruit fly (Edel and Kudla 2015). Homologues for this seem to be absent in all embryophyte species, however they are present in *Klebsormidium flaccidum*; a charophyte which is a sister group to land plants. Depolarisation activated channels are present in plant species and seem to show activity in response to environmental stresses such as osmotic pressure and microbe induction of depolarisation (Lhuissier et al. 2001; Okazaki et al. 2002). A similar group, hyperpolarisation-activated Ca^{2+} permeable channels, respond to abscisic acid (ABA) and reactive oxygen species through cytoplasmic Ca^{2+} influx in guard cells and root hairs of higher plants however the ion channels responsible have not been identified (McAinsh et al. 1990; Pei et al. 2000; Murata et al. 2001). While these two ion channels groups are not the same as VDCCs they do cover the same activation criteria and respond by allowing local increases in calcium ions thereby occupying a similar function with regards to calcium signalling in plants.

Mechanosensitive Ion Channels

In bacteria, mechanosensitive ion channels, or mechanosensitive channels, or small conductance channels (MscS), are involved with the movement of cations including calcium. Through the control of solutes, bacteria osmoregulate the cell and similar mechanisms seem to be found in plastids of *Arabidopsis thaliana* to control plastid size and shape (Martinac 2004; Haswell and Meyerowitz 2006; Veley et al. 2012). As plastids commonly act as intracellular stores of calcium it can be supposed that these channels may be involved in the efflux of calcium at those membranes. However, it is still unknown if these channels control calcium ion cytosolic influx in response to mechanical stimuli in plants (Trewavas and Knight 1994; Knight et al. 1991; Pleith and Trewavas 2002; Toyota et al. 2008). Doubt on the function of these channels in calcium signalling has increased with the discovery that at least two of these channels that are anion specific in *Arabidopsis thaliana* (Haswell et al. 2008; Maksaev and Haswell 2012). However, a related group of mechano-stimulated channels described as reduced hyperosmolarity-induced $[Ca^{2+}]^{i}$ increase channels (OSCAs) have been described independently by two groups as calcium specific channels. In *A. thaliana* OSCA mutants had impaired osmotic stress response in roots and guard cells, and involvement in innate immunity in guard cells providing a specific calcium channel that responds to processes known to trigger calcium signals (Yuan et al. 2014; Thor et al. 2020). New perspectives have arisen on mechanosensitive channels and their control of calcium in the mitochondria. MICU (Mitochondrial Calcium Uniporter) has become an increasingly studied model calcium ion channel. This channel is responsible for uptake of calcium into the mitochondria and the calcium signatures that occur in the mitochondria. Auxin and ATP triggered calcium responses occurred more rapidly and reached a higher maximal calcium concentration in *micu* mutants implying involvement in calcium efflux (Wagner et al. 2015; Wagner et al. 2016).

Two-Pore Channels

One type of channel known to function in animals is the two-pore calcium channels (TPC). These channels are similar to the VDCCs in that they are voltage gated channels. In animals these channels are thought to be involved in calcium-induced calcium release from the lysosome through the nicotinic acid adenine dinucleotide phosphate receptors (Brailoiu et al. 2009; Calcraft et al. 2009). However, these channels may not be specifically releasing calcium (Wang et al. 2012; Cang et al. 2013; Morgan and Galione 2014). In plants, TPCs are localised to the tonoplast membrane; the membrane of the vacuole. TCP1 in *A. thaliana* controls slow vacuolar current and influences the speed of salt-induced calcium peaks when knocked-out or overexpressed (Peiter et al. 2005a; Horecka and Chu 2017). However recent reports also suggest TCP1 may not be Ca^{2+} specific (Peiter et al. 2005a; Peiter et al. 2005b; Ranf et al. 2008; Beyhl et al. 2009; Ward and Schroeder 1994; Stephan and Schroeder 2014; Vincent et al. 2017).

Cyclic Nucleotide Gated Channels

Cyclic nucleotide gated channels are another channel that may non-specifically transport calcium ions in animals. These channels activate in response to cyclic nucleotides and are further regulated by calmodulin. CNGCs are found in prokaryotes as well as animals and plants and therefore represent a potentially ancient group of cation channels (Kuo et al. 2007; Nimigean et al. 2004; Hua et al. 2003b; Ali et al. 2006; Talke et al. 2003). However, a number of these channels have been demonstrated to show specific calcium ion release in plants, specifically *A. thaliana* CNGC 2/4/5/6/18 out of the 20 known (Zhou et al. 2014; Wang et al. 2013; Gao et al. 2014). These channels are known to be responsive to pathogens, abiotic stress and polar growth which provides a link between the CNGCs and potential calcium signals. Alongside this CNGC15-a/b/c are all involved in symbiotic interactions between *Rhizobium* and *Medicago truncatula* (Charpentier et al. 2016).

1.1.2 Calcium Signals are Encoded as Oscillations, Transients or Peaks

As the concentration of calcium in the cytosol is so much lower than that of each of the calcium stores or outside the cells, the opening of channels allows for a calcium concentration spike. The spike may repeat more than once hence the term calcium oscillation. Differential calcium ion oscillations can be precisely controlled to encode different responses to changes in conditions (stress responses) or development through temporal and spatial restriction of the signal (Fig.1.2). Ion channels are located in particular subcellular localisations and tissues which is known as a spatial restriction. Ion channels can also be expressed at particular points of a cell cycle or development of an organisms which can be considered a temporal restriction. Temporal restriction can also refer to activity of the ion channel generating the influx and efflux dynamics of a calcium signal. The membrane of the endoplasmic reticulum is a common site of calcium osciliations in animals, while in plants, the tonoplast membrane of the vacuole is also common. These two organelles in particular function as calcium storage compartments. Different calcium signals are referred to as calcium signatures and vary in the calcium ion concentration change, the duration and the frequency of calcium peaks. Each of these signatures may be bound by and activate subsets of multiple calcium decoding proteins or individual calcium decoding proteins (Zhu et al. (2007); Yu et al. (2006)).

Knowledge of calcium ion channels tells us that these calcium signatures take place and encodes the signal in response to particular stimuli. However, it does not tell us how the calcium signature can be decoded by downstream proteins to result in an organism response. Responses to calcium signatures take place by the activation of calcium decoding proteins which are responsive to different calcium signals. Using NFkB, JNK and NFAT as calcium responsive proteins the characteristics of calcium signals and the differential activation of calcium decoders by these signals was demonstrated. Differential activity of these transcription factors was shown in response to Ca^{2+} concentration, and duration, in single transients (Dolmetsch et al.

1997). In this study JNK, NFkB and NFAT were shown to all activate at high calcium concentration (1301±46 nM) when cells were forced to release calcium from their internal stores using ionomycin. EGTA was used to chelate calcium to abolish continuation of the calcium signal and it demonstrated that while NFkB and JNK remain activated after the signal, NFAT was quickly inactivated without a sustained signal. However, when the high calcium concentration transient was not generated but a low concentration plateau was sustained NFAT was activated and NFkB and JNK were not. This shows that different concentrations of calcium are required to activate different calcium responsive proteins, and that different durations are required to keep proteins activated (Dolmetsch et al. 1997). To further this study, a system to simulate calcium oscillations was set up by the flooding of cells with calcium and then EGTA. The activation of NFkB and NFAT was then compared. Different oscillations could not demonstrate differences in activity when calcium concentration was greater than 300 nM for these proteins, however at 275 nM the activity of these proteins could be distinguished. As shown before, NFkB persists after an activation signal so the protein was demonstrated to remain active even with a period greater than 1800 s however NFAT required oscillations every 400 s or less (Dolmetsch et al. 1998). Combined, these studies demonstrate that different proteins are responsive to different amplitudes of calcium as well as duration and frequency of calcium peaks.

In plants, guard cells have functioned as a good model system for calcium influx peaks controlling specific responses, through specific pathways regardless of the function of the ion channel that typically encodes the calcium peak. The guard cells respond to a range of signals such as reactive oxygen signalling, ABA response and external calcium addition by generating differential calcium signatures which results in stomatal closure (Allen et al. 2000). In studies on the *detiolated3* (*det3*) mutant, a putative V-ATPase, external calcium resulted in prolonged duration calcium signals that did not oscillate unlike the wild type response and stomatal closure was abolished (McAinsh et al. (1990); Brandt et al. (2015); Allen et al. (2000)). This

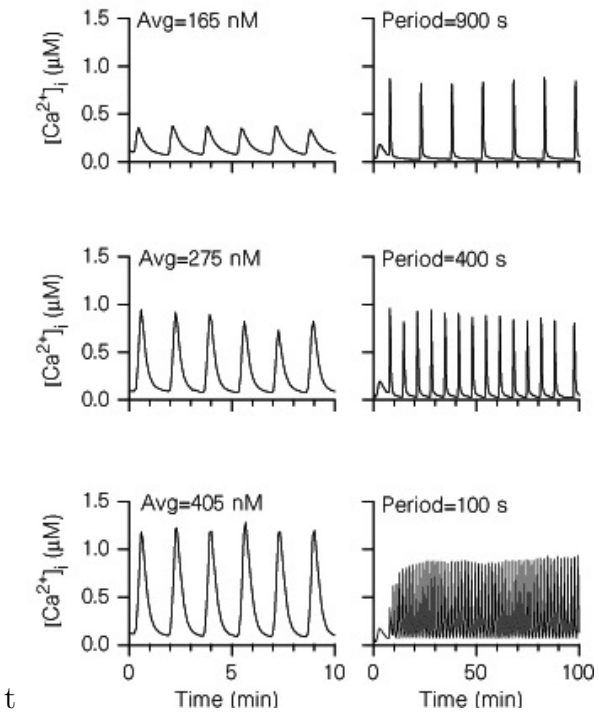


Figure 1.2: Different calcium signatures to trigger calcium decoder activity. Examples of different amplitudes of calcium signal with a constant frequency with concentration averaged in nM (left column). Different frequencies of calcium signal (right column) with frequency defined as period in seconds (Dolmetsch et al. 1998)

mutant had impaired sequestration of calcium which stopped oscillations but only in response to external calcium. The calcium signal was the same as wildtype in *det3* mutants when exposed to cold or abscisic acid. Furthermore, when *det3* mutants were forced to replicate wild type calcium signatures through the addition of calcium ion buffers and washing with buffers containing no calcium, stomatal closure was restored (Allen et al. 2000). Therefore *det3* demonstrates that pathways have specific signatures regardless of ion channel generating the calcium signature to result in the typical response. Similarly, to this mutant *gca2* also encodes an ion channel that is responsible for long term stomatal closure through the generation of a specific calcium signal (Allen et al. 2001). Three high concentration calcium oscillations spaced ten minutes apart close stomata by 65% which continued closure beyond an hour. If the oscillation is more or less frequent or of a lower amplitude the amount of closure is reduced, or the closure is not continued beyond an hour. In *gca2* the pattern of calcium oscillations has a higher frequency which stops long term

closure of stomata in response to ABA or reactive oxygen species. However typical responses can be restored in the *gca2* mutant by application of external calcium to generate the wildtype calcium signal (Allen et al. 2001). These two mutants demonstrate that the plant remains responsive to set calcium signals in predictable ways and therefore the decoding machinery is not affected by mutations in the encoding machinery if the same calcium signature can be produced. Additionally, they also demonstrate that very specific signals such as specific latency between oscillations results in specific responses such as long term stomatal closure.

1.1.3 Calcium Decoding Proteins Respond to Calcium Signals

As described before, calcium signatures vary in time and space through tissue expression, subcellular localisation, and so do the decoders in terms of specific expression in specific cell types or in response to specific stimuli. These decoders then activate in response to specific calcium signatures, usually through conformational change upon binding calcium followed either by protein-protein interactions or phosphorylation through intrinsic kinase activity. The calcium decoder proteins can then regulate activity of targets, which can be transcription factors resulting in sweeping changes to gene expression, or ion channels to extrude toxins such as Na^+ extrusion by Na^+/H^+ antiporters. Stimulation of three diverse types of calcium signature resulted in up or downregulation of a range of genes in the transcriptome with 61.1% or more unique to each signal (Whalley and Knight 2013). The changes in transcriptome in response to these calcium signatures has allowed for the characterisation of several promoter sequences linked to these signals including the abscisic acid responsive-element (ABRE), C-Repeat/Drought-Responsive Element, Site II and CAM box (Whalley et al. 2011).

Calcium decoders detect calcium fluctuations by binding calcium through intrinsic helix-loop-helix domains, known as EF hands, and fall broadly into sensor-relays and sensor-responders types (fig 1.3. Sensor-responders, such as Calcium Depen-

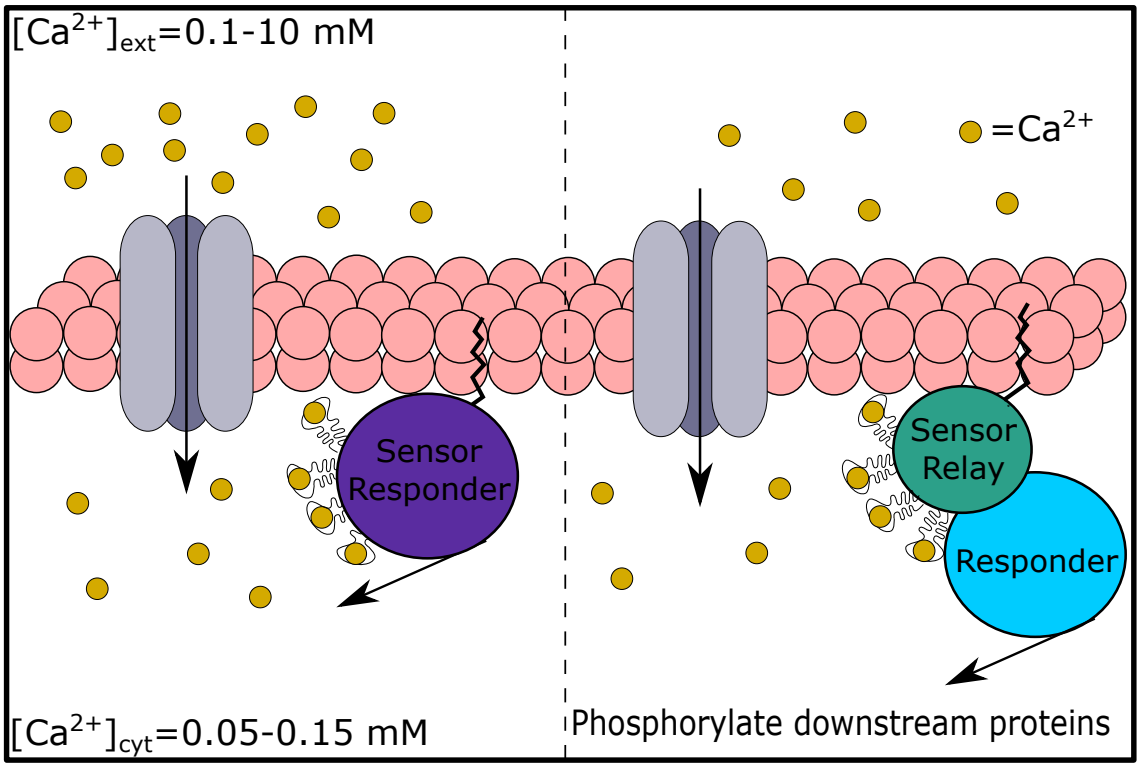


Figure 1.3: Models of a sensor-responder type calcium decoder and a sensor-relay type calcium decoder. Sensor-responder type calcium decoders detect a calcium signal and have direct activity on protein targets such as ion channels to modulate their activity. Sensor-responder calcium decoders detect a calcium signal and will interact with another protein and will change in conformation in response to the signal which will release inhibition on the interacting protein which will modulate activity of downstream proteins.

dent Protein Kinases (CDPKs), detect calcium changes through EF hands and then phosphorylate downstream proteins directly with intrinsic kinase domains. Sensor-relays, such as Calcineurin B-Like proteins (CBLs) and Calmodulins (CaMs) have no intrinsic kinase activity but upon binding of calcium undergo a conformational change which allows them to interact with downstream partners to regulate the activity of other signalling proteins (Kolukisaoglu et al. 2004). Calcium decoders such as calmodulin can regulate both transcription factor activity and the channels that encode other calcium signals such as CNGCs. Comparing the structure of these decoding proteins is elucidating some of their regulatory mechanisms and providing new understanding about how they function.

1.2 Comparative Structure of the Calcium Signal Decoding Proteins

1.2.1 Calcium Decoders have Different Calcium Affinities Due to EF hand Structure

Calcium decoding proteins are identifiable by their EF hand domains responsible for Ca^{2+} coordination and these residues bind calcium ion directly through a calcium binding loop (Sánchez-Barrena et al. 2013). Canonical EF hands have a consensus helix-loop-helix motif consisting of around 30 amino acids (Sánchez-Barrena et al. 2013). The loop structure is the most divergent in sequence typically consisting of 12 amino acids and changes in this sequence seem to account for most of the difference in calcium binding affinity (Albrecht et al. 2001) with a consensus of DxDxDGxxDxxE (Kolukisaoglu et al. 2004; Batistič et al. 2010; Li et al. 2009). However, the loop sequence can vary from 10 and 20 amino acids (Gentry et al. 2005; Silva et al. 1995) Positions 1 (X), 3 (Y), 5(Z), 7 (-X), 9 (-Y) and 12 (-Z) of the 12 amino acid canonical EF hand sequence are directly responsible for the coordination of calcium. The coordination positions of the loop link to calcium through 7 oxygen atoms into a pentagonal bipyramid with the Glu in the -Z position contributing two oxygens (Silva et al. 1995). However, a vitamin-D dependent calcium binding protein has an N-terminal EF hand which differs from the consensus in that all calcium coordinating ligands are main chain carbonyl oxygen atoms without the typical Gly in position 4 of the canonical loop (Szebenyi and Moffat 1986). Comparison of longer than canonical loops demonstrate that their calcium binding affinity is more tolerant to insertions in the N-terminus such as the 14 amino acid EF hand in CBLs (Albrecht et al. 2001).Therefore the EF hand loop structure can bind calcium through a wide range of mechanisms either through typical loop structures or completely different coordination chemistries.

X-ray crystallography of a range of calcium binding proteins has helped to eluci-

date the mechanism by which calcium is bound and the signal is transduced through conformational change. The conformational change that results is predominantly in the two helices either side of the loop due to the movement of a hydrophobic residue (Ile, Leu or Val) which flanks the -Y Asp, when the calcium is coordinated to the Glu. A number of EF hands have Asp instead of Glu in the -Z position which results in reduced specificity for calcium and increases affinity for Mg^{2+} binding (Silva et al. 1995). However crystal structure has demonstrated that the binding of magnesium does not result in the conformational change associated with calcium binding. The two-step hypothesis of calcium binding is that the N-terminus is tolerant to mutation because it is only required for binding of divalent cations, the specificity in EF hands for calcium comes from the -Z Glu residue and position 4 Gly for flexibility of the loop (Silva et al. 1995). However, it is the specific movement of this -Z position amino acid to coordinate calcium that results in the conformational change required for activation. Therefore while magnesium binding is possible it does not result in conformational change required for calcium signalling, but it cannot be ruled out that it plays a modulating role for the signal through competitive binding.

The conformational change of one EF hand does not only affect that single EF hand but changes the calcium environment and tertiary structure of the protein and therefore can affect other EF hand domains on the same protein. Calmodulin contains four EF hands with each hand capable of binding calcium and when the proteins is lysed with trypsin it forms two domains of pairs of EF hand that are also functional in binding calcium (Drabikowski et al. 1984) but single EF hands are not functional and therefore protein hands are typically considered in pairs (Drabikowski et al. 1984). Pairs of EF hands coordinate a single calcium to one EF hand which results in a conformational change that can affect the affinity of its pairmate for calcium which is known as EF hand cooperativity. EF hands cooperativity is brought about due to the multiple hydrophobic contacts between the helices of the pairs and hydrogen bonds directly between the calcium binding loops forming a short antiparallel B-sheet. The CaM(Q41C/K75C) mutant introduces a disulphide bond

between helices B and C (4.14A) that is similar in distance to native conformation distance (7.14A). The helices of these two EF hands are usually mobile with respect to each other but the covalent bond locks the structure into an inactive conformation (Grabarek 2005). However, the N-terminal part of the EF hand which is responsible for initial calcium binding, can still bind calcium in the Cam41/75 mutant but the position12 Glu does not move after coordinating calcium. Cam41/75 had no cooperativity in EF hands suggesting it is the physical movement of the helices with respect to each other that governs cooperativity (Grabarek 2005).

Over 200 proteins have been identified in *Arabidopsis thaliana* as containing domains that sense cytosolic calcium (Day et al. 2002; Reddy and Reddy 2004). Many calcium decoders such as CaMs and CDPKs contain four EF-hand domains. These domains do seem to function in cooperative pairs as previously described and each EF hand in the pair affects the calcium binding affinity of the other. In CDPKs the C-terminal EF hands are high affinity calcium coordinators that bind calcium at basal concentration and stabilise the inactive conformation (Wernimont et al. 2010). The two N-terminal EF hands are low affinity, bind calcium during the specific calcium signal and release CDPK inhibition (Christodoulou et al. 2004). For example, in AtCPK-1 the C-lobe has a Kd of 5.6 nM at 20 mM KCl which can bind calcium at basal concentrations and the N-lobe has a Kd of 110 nM at 20 mM KCl which is above basal calcium (Christodoulou et al. 2004). However, four canonical EF are not compulsory as CBLs contain three canonical EF hands. The first EF hand in CBLs contains an insertion of two amino acids between position X and Y which can still bind calcium (Kolukisaoglu et al. 2004; Sánchez-Barrena et al. 2013; Negae et al. 2003). Similarly, Calcium- and Calmodulin-dependent protein kinases (CCaMKs) contain a visinin-like domain with four EF hand domains; however only three are functional (Ramachandiran et al. 1997; Swainsbury et al. 2012).

1.2.2 Calmodulin as a Calcium Sensor-relay Protein

Calmodulins (CaMs) are ubiquitous in eukaryotic cells and are some of the most well studied calcium signalling proteins (Yang and Poovaiah 2003). *Arabidopsis thaliana* has four distinct CaM isoforms encoded by seven genes (Zhu et al. 2015; McCormack and Braam 2003). CaM has no intrinsic activity itself, but can regulate a wide range of protein targets including transcriptional regulators, signal proteins and channels (Yang and Poovaiah 2003; Popescu et al. 2007; Hua et al. 2003b; DeFalco et al. 2016). Some calcium-dependent proteins are also regulated by CaMs such as CCaMK, CPK28 in *A. thaliana*, and Calmodulin-dependent Kinases (CaMKs) commonly found in animals (Ramachandiran et al. 1997; Bender et al. 2017; Watillon et al. 1995). CaM proteins contain four EF hands with a lot of conservation between each isoform which has made it difficult to determine specificity in their protein targets and hence for which proteins they activate or inhibit activity. Upon calcium binding CaM proteins go through conformational changes to reveal a hydrophobic region which acts as a binding domain. Therefore all of their activity is thought to be due to protein-protein interactions and the changes in tertiary structure or physical occlusion that interaction causes in target proteins.

1.2.3 Calcineurin B-Like Proteins as Specific Calcium Sensor-relay Proteins

Calcineurin B-Like proteins (CBLs) are similar in structure to CaMs. Their calcium binding domains have four functional EF hand with the first EF hand consisting of a 14 amino acid consensus sequence (Negae et al. 2003; Sánchez-Barrena et al. 2013). Spatial specificity is common in these proteins due to N-terminal variable domains that target the proteins to subcellular locations. N-terminal myristoylation of a glycine residue and S-acylation of the adjacent cysteine residue in the consensus MGCXXS/T domain confers plasma membrane targeting (Batistič et al. 2010). Tonoplast targeting is conferred by S-acylation of four cysteine residues in the first 22 amino acids of *A. thaliana* CBLs (Batistič et al. 2012). Uniquely, *A. thaliana*

CBL10 contains a transmembrane domain for tonoplast targeting (Kim et al. 2007). However, this may not be as specific as first proposed as it seems that under different conditions this protein may localise to different locations through unknown mechanisms which may allow AtCBL10 to respond to intracellular spatially separated signals under different conditions (Batistič et al. 2010; Quan et al. 2007; Ren et al. 2013). CBL sensor-relays interact specifically with downstream CBL-interacting protein kinases (CIPKs). Different CBLs can interact with different CIPKs under different conditions. With ten CBLs and twenty-six CIPKs resulting in 260 potential combinations in *A. thaliana* ascribing functions has been difficult (Waadt et al. 2008; Batistič et al. 2010).

1.2.4 CBL-Interacting Protein Kinases as Mediators of Calcium Signals

There are multiple types of calcium decoders with an intrinsic serine/threonine kinase domain for downstream signalling through phosphorylation. Some calcium decoding proteins have intrinsic calcium binding domains (Fig.1.4). However some are kinases without intrinsic calcium binding domains but interact with calcium sensor proteins such as CBLs and CaMs and these calcium sensors respond to specific calcium signatures thereby conferring calcium specificity. Of these, CIPKs do not have an intrinsic calcium binding domain but interact with specific CBLs which detect calcium. For example, AtCBL4 specifically interacts with AtCIPK24 under a salt stimulus triggered calcium signal in *A. thaliana* roots but AtCIPK24 can interact with AtCBL10 in the shoots of *A. thaliana* (Kim et al. 2007). CIPKs have N-terminal kinase domains and a C-terminal inhibitory junction. On this junction there is a domain for interaction with CBLs known as a NAF domain (Fig.1.4; Albrecht et al. (2001)). Hydrophobic interaction with the NAF domain (named for based on conserved amino acid sequence) in CIPKs allows phosphorylation of the FPSF domain (also named based on amino acid sequence) of the CBL which is thought to be involved with the pairing of these proteins and substrate specificity

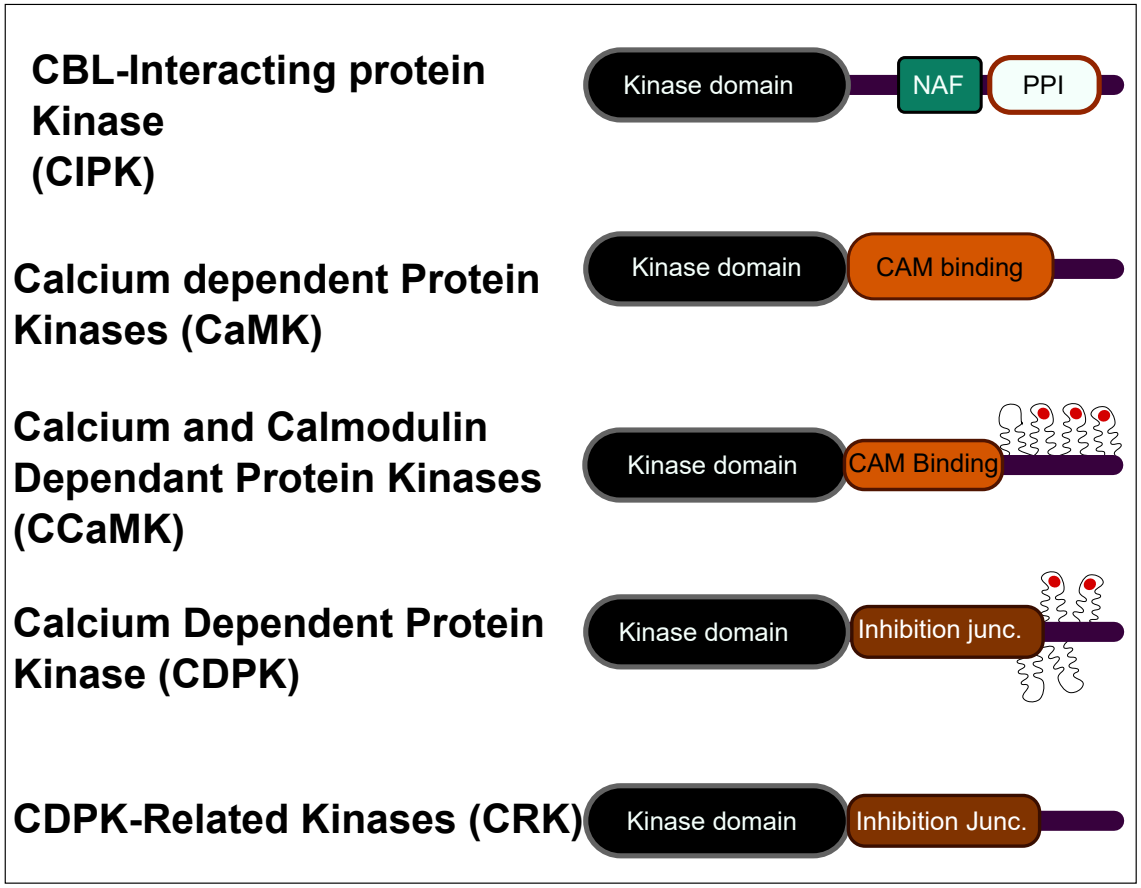


Figure 1.4: Models of a range of calcium decoding proteins to compare important domains for downstream interactions. CIPKs contain an intrinsic kinase domain, a NAF domain (named for key residues) for interaction with CBLs, and phosphatase interaction motif (PPI) . CAMKs are similar with a kinase domain and a CaM binding domain (CaM BD) for interaction with the CaM calcium sensor. CCaMKs have both of the CaMK domains with the addition of a Calcium binding domain with three functional EF hands. CDPKs have the same domains as a CCaMK with four functional EF hand domains, most cannot bind CaM. CRKs are similar to CDPKs but have a range of function EF hands, many have none.

(Hashimoto et al. 2012).

1.2.5 Calmodulin Dependent Kinases as Sensor Interacting Kinases

Similarly to the CIPKs, calmodulin dependent kinases (CaMKs) have a highly conserved serine/threonine kinase domain. However, some CaMKs also have an N-terminal variable domain before the kinase domain (Nagata et al. 2004). These proteins also do not contain EF hands and have a variable C-terminal region which

contains a CaM binding domain (Fig.1.4; Zhang and Lu (2003)). The CaMK family is more diversified in animals than in plants and only three CaMKs are present in *A. thaliana* (Nagata et al. 2004). Of the CaMK proteins there are four groups: CaMKI, CaMKII, CaMKIII, CaMKIV. Of these groups, CaMKIII have single specific substrates but the others have multiple downstream targets (Hook and Means 2001). CaMKI/II/IV groups are most closely related to plant calmodulin dependent kinases (Chen et al. 2017). Very few proteins that are similar to this group have been found in plants except three in *A. thaliana*, two in rice (*Oryza sativa*) and one in apple (Nagata et al. 2004; Watillon et al. 1992; Watillon et al. 1993; Watillon et al. 1995).

1.2.6 Calcium Calmodulin Dependent Protein Kinases Reveal a Potential Activation Model

CCaMK is a widely studied calcium decoding protein important in plant symbiosis events (Lévy et al. 2004). CCaMK is thought to have arisen from a fusion between a CaM and a CaMK to make the group CaM Fused Kinases (CFK; Chen et al. (2017)). As such these proteins contain the kinase domain, autoinhibitory junction and calmodulin binding site of a CaMK and a unique calcium binding domain. This calcium binding domain is a visinin-like domain with three functional EF hand motifs in CCaMK but varies in other CFK (Fig.1.4; Ramachandiran et al. (1997)). Studies of the activation of CCaMK have formed a general model that may be true for other CFKs. It is proposed that the pseudosubstrate domain of the autoinhibitory cleft occludes the catalytic domain of the kinase. This interaction is thought to be stabilised by the EF hand domains which bind calcium at basal concentration. The binding of the calcium allows for autophosphorylation in Thr-271 in the kinase domain to stabilise a hydrogen network thereby stabilising the inactive conformation. Interestingly, the Thr-271 phosphorylation also exposes adjacent amino acids to allow for CaM binding which is associated with CCaMK activation therefore calcium binding also 'primes' the CCaMK for subsequent activation by CaM (Fig. 1.5; Miller

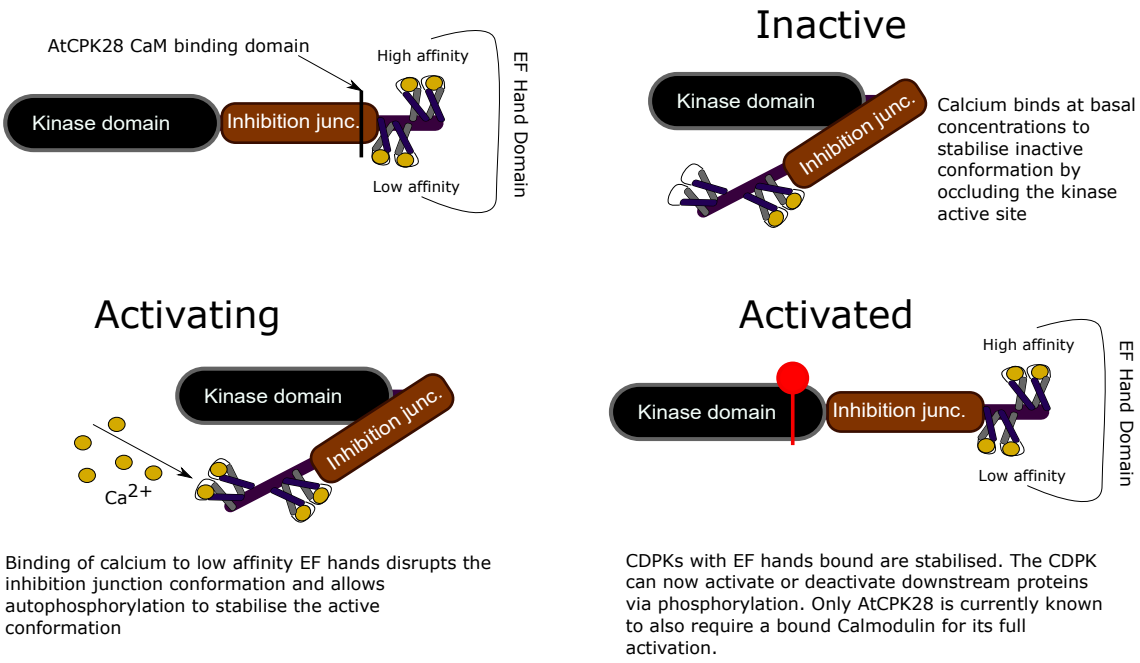


Figure 1.5: Structure and activation model for the CDPKs. Top panel shows the structure and main regions of a CDPK protein. Top right is the inactive CDPK with intramolecular forces holding it in a inactive conformation and calcium bound to two EF hands at basal concentration. Bottom left is an activating CDPK in response to a calcium oscillation. Calcium binds to the low affinity EF hands resulting in disruption of intramolecular forces in the autoinhibitory junction that allows for autophosphorylation. A threonine is represented by Thr but this residue(s) could be serine or tyrosine as well. Bottom right is an active CDPK with calcium bound to the four EF hands, the active kinase is exposed and capable of phosphorylating interacting proteins.

et al. (2013)).

1.2.7 Calcium Dependent Protein Kinases as a Widely Diversified Group in Plants

CDPKs are a widely diverse group of CFK proteins in plants with 34 proteins in *A. thaliana*. These proteins fall into four subgroups with some similarity in function between individual members but functional redundancy between the members have hampered characterisation. CDPKs have many similar domains to the CCaMKs in that they are also thought to be part of the CFK group evolved from the fusion of a CaM and CaMK. As such they are serine/threonine kinases with an autoinhibitory junction and calcium binding domain or CaM-like domain (CLD) (Fig.1.4).

Similarly to CCaMK, the autoinhibitory junction and CLD domains are involved in the stabilisation of inactive conformation at basal calcium concentrations (100 nM) by occluding the catalytic cleft (Harmon et al. 1994; Harper et al. 1994; Chandran et al. 2006; Christodoulou et al. 2004). However, unlike the CCaMKs these proteins have four functional EF hands, and most CDPKs cannot bind CaM. In the case of CDPKs, calcium binds to the C-terminal EF hand pair to stabilise the protein's inactive conformation and this stability is released upon binding of calcium to the N-terminal EF hand pair (Christodoulou et al. 2004; Wernimont et al. 2010). The N-terminal EF hand pair have a lower affinity for calcium thereby binding calcium at higher concentrations caused during calcium signalling events. Further specificity is added by a variable N-terminal domain that allows for subcellular localisation like that found in the CBLs therefore they can only respond to signals based at their subcellular localisation (Fig.1.5; Ito et al. (2010)).

1.2.8 The Study of Calcium Signalling is Hampered by Understudied Protein Families

In addition to CaMKs, many other protein families have been neglected in the study of calcium decoding in higher plants. CDPK-related proteins (CRKs) are an example that have a similar structure to CDPK but with degenerate EF hands, some can even bind CaM which makes them similar to CCaMKs (Wang et al. 2004). Two other groups of calcium decoders are known as Calmodulin binding proteins (CBKs) and 'Maize homologues of mammalian CaMK' (MCKs) and have no distinct difference from CaMKs found in animals but lack sufficient study to link them together. CBKs and MCKs are both very poorly characterised protein families in higher plants that contain a calmodulin binding domain and a kinase domain (Hua et al. 2003a; Lu et al. 1996; Wang et al. 2001). Two other neglected protein groups in plant studies are phosphoenolpyruvate carboxylase kinase (PPCKs) and PPCK-related kinases (PEPRKs) which seem to have formed from the fusion of CaM and a kinase, the same as CCaMKs and CDPKs, but have lost the autoinhibitory domain and CaM

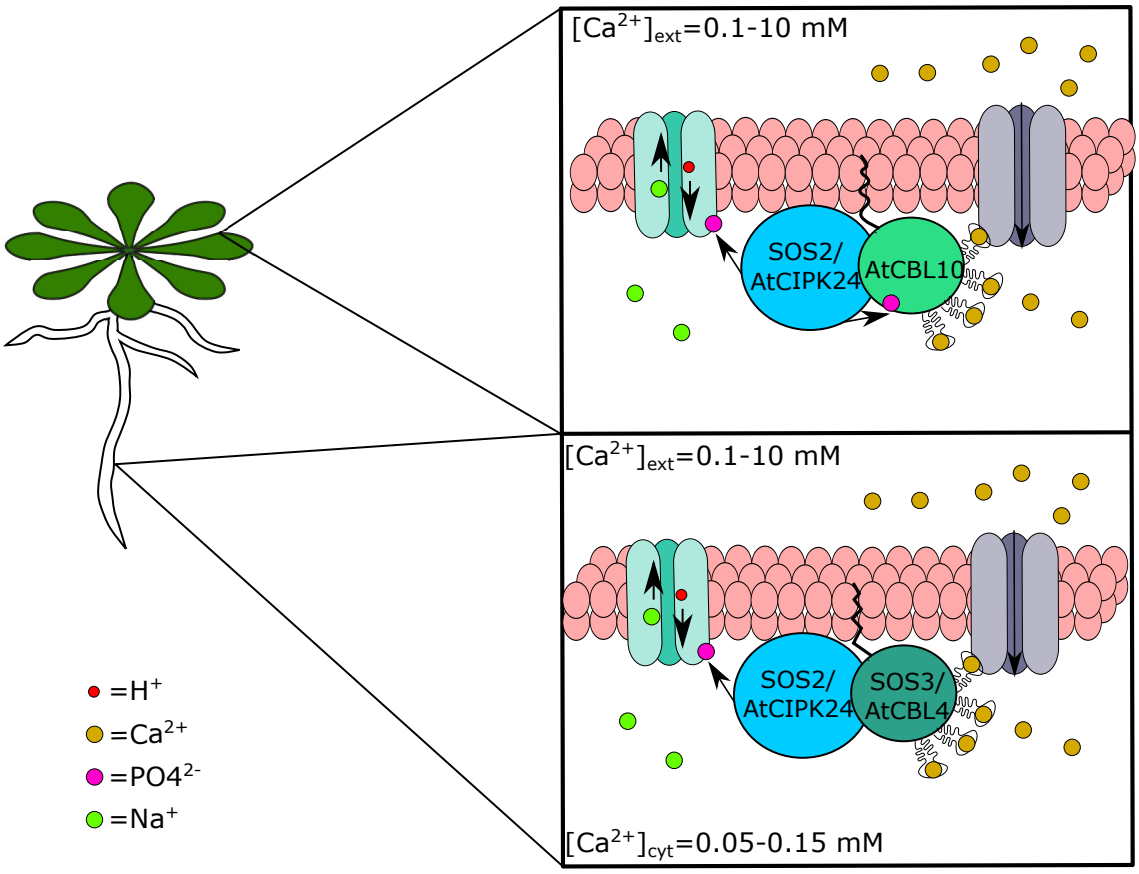


Figure 1.6: [Model of the Salt Overly Sensitive pathway from *A. thaliana*. In the shoot AtCBL10 detects a salt induced calcium signal and binds the calcium which activates SOS2/AtCIPK24 which phosphorylates AtCBL10 and SOS1 which will export Na^+ to reduce the toxic sodium effect. In the root SOS3/AtCBL4 detects a salt induced calcium signal and binds the calcium which activates SOS2/AtCIPK24 which phosphorylates SOS1 which will export Na^+ to reduce the toxic sodium effect and occlude sodium from entering the transpiration stream.

binding domain (Chen et al. 2017). It has also been shown that some of these PPCKs and PEPRKs have calcium independent activity which implies that they may have derived from but do not necessarily any longer function as calcium decoding proteins (Furumoto et al. 1996).

1.3 Calcium Decoding Proteins Respond to Stress Induced Calcium Signals

1.3.1 CBLs Respond to Salt Induced Calcium Signals

Calcium decoding proteins interact with downstream proteins to control their function. Thereby, this interaction can directly stimulate a signalling cascade to result in a noticeable activity such as a stress tolerance. *Salt-Overly-Sensitive* (*sos*) mutants in *A. thaliana* are an example of how CBLs and CIPKs can modulate a downstream ion channel in response to salt stress. Interest was developed in *sos1* when it was characterised to be sensitive to Na^+, Li^+ and K^+ and overexpression of *SOS1* was found to increase salt tolerance. SOS1 was then characterised as a 127 kDa protein with a 12 transmembrane domain that functioned as a cation channel based on similarity to bacterial and fungal channels. Weaker salt sensitivity phenotypes were found in *sos2* and *sos3* mutants. SOS2 was bioinformatically homologous SnRK3, now known as a CIPK, and SOS3 showed similarity to calcineurin and recoverin, and was one of the first characterised CBLs (Ji et al. 2013). The SOS pathway has now developed into a model system for CBL/CIPK complex activity in response to salt stress (Fig. 1.6, Ji et al. (2013)).

SOS1 expression is upregulated in high salt conditions and the protein product is a cation channel regulated by SOS2 (AtCIPK24) and SOS3 (AtCBL4) (Ji et al. 2013). Vesicle studies demonstrated conclusively that SOS1 was a Na^+/H^+ exchanger as it depleted electrochemical gradients to transport Na^+. It was also shown that activity of SOS1 was increased when exposed to auto-active SOS2 and therefore that the SOS2/3 regulation was not just at an expression level but at direct regulation of the protein channel. Alongside this, depletion of the electrochemical gradient also elucidated that the channels for Na^+ transport also require H^+ exchange activity (Qiu et al. 2002). Reconstitution of the pathway in yeast demonstrated that SOS2 (AtCIPK24) and SOS3 (AtCBL4) were required for increased SOS1 activity, and that SOS2 (AtCIPK24) and SOS3 (AtCBL4) together increase activation more

than auto-active SOS2 (AtCIPK24) can by itself. The channels were also shown to be specific to Na^+. In yeast, SOS3 (AtCBL4) activated in response to calcium stress signals and formed a complex with SOS2 (AtCIPK24) to localise it to the plasma membrane. Auto-active SOS2 (AtCIPK24) was not localised to the membrane which is the location of the SOS1 channel resulting in a less efficient activation of SOS1 (Quintero et al. 2002). Further proteomics was used to identify that the C-terminal tail domain of SOS1 was required for the SOS2/3 regulation and that deletion of this domain lead to reduced activity but also independence from SOS2/3. Further analysis characterised this autoinhibitory C-terminal tail to find a consensus [V,I]VR[V,I]D<u>S</u>PS motif which is phosphorylated on underline serine residue. This established the paradigm that a CBL (such as SOS3) forms a complex with CIPKs (such as SOS2) to result in subcellular localisation of downstream targets as well as to confer calcium sensitivity (Fig. 1.6). CIPKs are responsible for the phosphorylation to up or downregulate the activity of the downstream protein, in this case a cation channel (Quintero et al. 2011).

However, there is further complexity in this model as seen by the fact that *SOS1* also has higher expression in high salt conditions and the phenotype of *sos2/3* is weaker than that of *sos1* (Shi et al. 2002). Components of the SOS2/3 regulatory system are also controlled by expression. *SOS3* expression is predominantly in the roots whereas *SOS1* and *SOS2* are expressed in all parts of the plant. SOS1 loads Na^+ into xylem in the shoot but it cannot be regulated by SOS3 to do this. A study found that AtCBL10 characterised as SOS3-Like Calcium Binding Protein 8 (SCaBP8) was expressed mostly in the shoot and interacts with SOS2 (AtCIPK24) to localise it to the plasma membrane in the same way SOS3 (AtCBL4) does in the roots. This demonstrates a tissue specific change in interaction partner for SOS2 (AtCIPK24) (Quan et al. 2007). There are some differences in the interaction between SOS3-SOS2 and AtCBL10-SOS2 as AtCBL10 requires phosphorylation by its interacting protein kinase and does not have the MGCxxS motif to localise it to the plasma membrane but instead localises to the tonoplast usually and changes

localisation to the plasma membrane under and unknown mechanism (Fig. 1.6).

CBL and CIPK interactions do not only extrude ions from the plants but can actually be involved in uptake under nutrient difficiency or nutrient abundant conditions. Also partners of CBL and CIPK can interact with more than one ion channel downstream which is not represented in the SOS pathway. For example, the AtCBL1/9 both interact with AtCIPK23 to regulate positively potassium uptake by phosphorylating and increase activity of high-affinity K^+ transporter 5 (HAK5) in low external K^+ conditions (<10 µM). AtCBL1/9 can also interact with AtCIPK23 to increase activity of the low affinity Arabidopsis K^+ transporter 1 (AKT1) in high external potassium conditions (>500 µM) (Li et al. 2006; Cheong et al. 2007; Ragel et al. 2015). The exact mechanism for how the CBL-CIPK complex recognises high or low potassium and targets the correct channel is not elucidated, it could be calcium signature specific, phosphorylation pattern specific, or protein partner specific.

The regulation of protein activity of SOS1 is also more complex than just regulation by the SOS pathways as SOS1 is also phosphorylated by MPK6 which is implicated in increased salt tolerance (Yu et al. 2010). The exact site of MPK6 phosphorylation is unknown but the site of SOS2 phosphorylation also overlaps with a putative MAPK phosphorylation site. The putative MAPK phosphorylation site is a serine two amino acids N-terminal of the CIPK phosphorylation site. The putative MAPK target serine residue is key to the interaction between SOS2 (AtCIPK24) and SOS1 and alanine mutation of this site abolished phosphorylation by SOS2 (AtCIPK24) (Quintero et al. 2011). However, as this mutation abolished phosphorylation by SOS2 (AtCIPK24) it may be that either MPK6 or SOS2 can phosphorylate the S1136 or S1138 site respectively to enhance salt tolerance which may explain why *sos2* sensitivity is less than that of *sos1*. SOS2 (AtCIPK24) is negatively regulated itself though other processes adding further complexity to the SOS pathway. It was demonstrated that 14-3-3 proteins bind and reduce activity of SOS2 (AtCIPK24) if the protein is already phosphorylated (Zhou et al. 2014). As

this phosphorylation only takes place when SOS2 (AtCIPK24) is interacting with a CBL partner which becomes activated by a calcium change, the 14-3-3 protein inhibition is a direct inhibitor of SOS2/3 regulation of SOS1. Therefore much of the surrounding regulation of the SOS pathway still needs characterising to understand all of the mechanics of how CBLs and CIPKs are involved in salt tolerance responses and what other proteins are implicated in the tolerance conferred.

1.3.2 CDPKs Respond to Help Assimilate Nitrate in Abundant Conditions

As previously mentioned CDPKs respond to many types of stress including triggering innate immunity in response to biotic stress, salt tolerance, drought tolerance (Bender et al. 2017; Fu et al. 2013; Brandt et al. 2015). However, AtCPKs can also respond to abundant nutrition such as high nitrate to coordinate uptake and subsequent growth networks. Nitrate is a primary assimilate for many plants and can be a limiting factor for growth. Due to this nitrate fertilisers are used for crop growth. Much of this nitrate is produced from ammonium that is generated by the Haber-Bosch process which requires high energy and represents 1% of the world's energy consumption. It has recently been shown that CDPKs play a significant role in nitrate signalling in plants (Liu et al. 2017). Calcium signals have been detected in the cytosol and nucleus of plants treated with nitrate but not in response to other nitrogen assimilates such as ammonium and glutamine. Expression of the *NIR* transcription factor is upregulated by nitrate addition. Fusing the NIR promoter with a luciferase reporter gene results in bioluminescence which can be directly used to monitor signalling in response to nitrate. Constitutively active CDPK mutants were generated and an autoactive NIR-luciferase response was detected with subgroup three CDPKs (CPK-7/8/32/10/30/13). Single knockout mutants were generated for each of these proteins and little change in nitrate response was detected due to functional redundancy between them. Therefore, higher order mutants were generated with the highest responders to increased nitrate including AtCPK10, AtCPK30 and

AtCPK32. However, the double knockout of CPK10/30 were embryonic lethal which meant it could not be assessed for involvement in nitrate uptake. To circumvent this problem a point mutation was generated in the ATP gatekeeper residue of CPK10 (M141G). This point mutation mimicked CaMKIIα which is reversibly inhibited by 3MBiP (1-isopropyl-3-(3-methylbenzyl)-1*H*-pyrazolo[3,4-*d*]pyrimidin-4-amine), and made it possible to inhibit CPK10 with 3MBiP. Due to this reversible inhibition, the triple mutant CPK10(M141G)/30-/32- was no longer embryonic lethal and the active CPK10 could be inhibited by 3MBiP. Quantitative PCR was used to detect changes in gene expression in the triple mutant, and a range of factors including those involved in nitrate assimilation and transport were not upregulated in the triple *AtCPK10/30/32* mutant but were in the wild type. Interestingly, shoot development, glycolysis and oxidative pentose pathway factors were upregulated in the wild type and all are directly involved in uptake of nitrate and other nitrogen containing products such as ammonium and glutamine, even though ammonium and glutamine did not effect AtCPK10/30/32 activity. A growth sensitivity study was also used to measure shoot development and nitrate, and it was demonstrated that both calcium and these CDPK proteins were required for signalling in response to nitrate. GFP fusions also showed that all three of these CDPK proteins translocate to the nucleus in response to nitrate stress thereby demonstrating a localisation effect and method for which they can act upon transcription factors (Liu et al. 2017).

1.3.3 Functional Redundancy Hampers Studies of Calcium Decoding Proteins

Unravelling the calcium decoding network has been hindered in higher plants by functional redundancy. The functional redundancy can be exemplified in the nitrate study (Liu et al. 2017) but other studies have also had to make higher order mutants to detect phenotype (Boudsocq et al. 2010). The overlapping function of the CDPKs, and other calcium decoders such as CBLs, means that individual gene knockouts may not give rise to a phenotype which makes it more difficult

to decipher the function of each protein and progressively more complicated and situational methods have been used to actually ascribe functions to proteins such as making AtCPK10(M141G) mutant to make it chemically inhibitible. Therefore, early-diverging land plants may be useful in decoding calcium signals. Early-diverging land plants have not gone through as many genome duplication events and therefore have a reduced genome size and likely a reduced calcium signalling repertoire. This will allow for reduced functional redundancy so single knockouts can lead to phenotypes which will help to characterise function. The understanding derived from this will be applicable to later-diverging plants as it is likely that groups of calcium signalling proteins will correspond to similar functions in calcium decoder subgroups in later-diverging plants. Some groups of calcium decoder proteins have conserved functions with closely related calcium decoders. For example AtCIPK24 is most closely related to AtCIPK8 and both can activate SOS1 in response to salt stimulus calcium signatures through interaction with AtCBL10 in *A. thaliana* shoots (Yin et al. 2019). It is likely that this functional redundancy occured due to whole genome duplication and therefore a single CIPK may be present in a plant which has not gone through whole genome duplication which can represent the early function of this clade. However some members of this clade could had later derived new function through neofunctionalisation post whole genome duplications which may mean that some functions may not be present in sister clades like the bryophytes. Regardless, characterising calcium decoders in early-diverging plants may provide insight into early functions of different calcium decoder subgroups which will help to propose functions for currently uncharacterised calcium decoders in later-diverging plants.

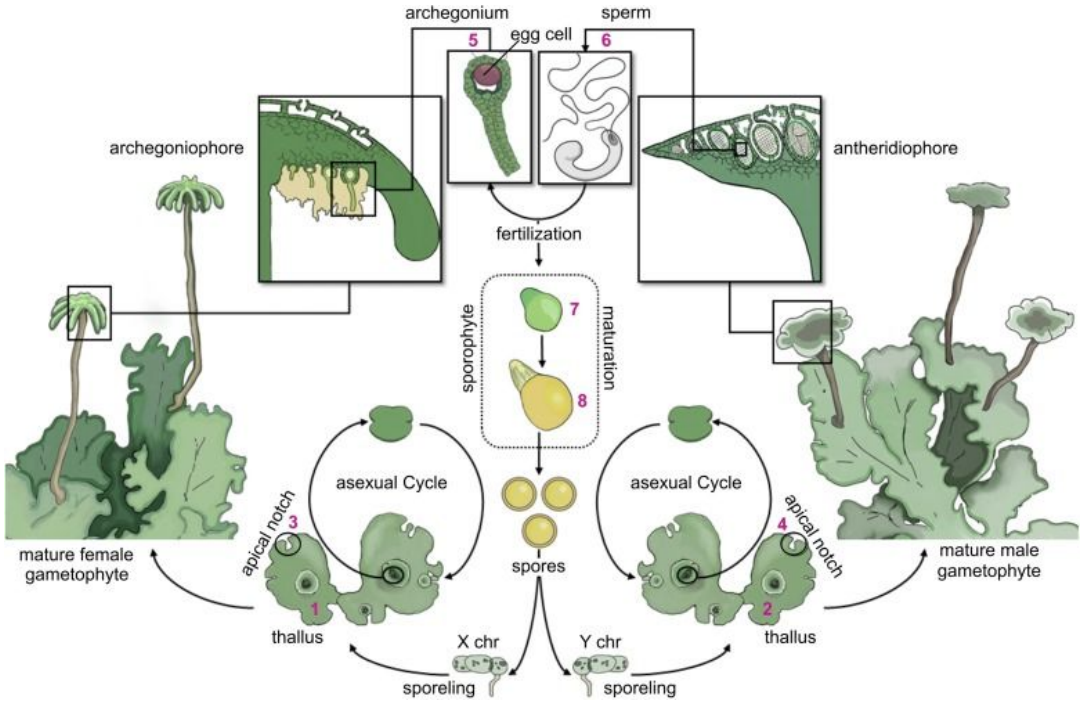


Figure 1.7: Illustration of the *Marchantia polymorpha* lifecycle including both sexual and asexual cycles. Magenta numbers represent particular tissues. 1/2 are male and female thallus respectively, apical notch of male and female gametophyte (3/4), archegonia (5), antherozoids (6), early sporophyte (7), and late sporophyte (8). Early and late sporophyte (7/8) phases and sores are the diploid phases of the haploid dominant plant (Schmid et al. 2018)

1.4 *Marchantia polymorpha* as a Model Basal Land-plant

M. polymorpha is a phenotypically well characterised liverwort with a unique position at the base of the green lineage as one of the first land plants (Bowman et al. 2007). Due to its lineage it has not gone through the genome duplication events which are thought to be responsible for the functional redundancy that has made unravelling the function of the calcium decoders so difficult in higher plants. *M. polymorpha* also has a dominant haploid life phase which opens the potential for a range of experiments which are not possible in higher plants (Fig.1.7).

1.4.1 The Life Phases of *M. polymorpha* Make it a Useful Organism for Study

M. polymorpha subsp. ruderalis is a liverwort and a dioecious basal land plant. Like many land plants it goes through an alternation of generations with a dominant phase of vegetative growth and a shorter phase of sexual growth and reproduction. Unlike flowering plants, *M. polymorpha* has a dominant haploid phase of vegetative growth where it has one copy of the genome. When the plant reproduces it forms a sexual structure called a gametophore. In the males the sexual structure is called the antheridiophore and it produces motile male gametes called atherozoids or spermatozoids (Fig.1.7). In the female, the gametophytes are called archegoniophores and have venters (cavities) which hold a single egg cell. The sperm cells will come into contact with the egg cells to produce diploid sporophytes in the archegonium. As the sporophyte grows in the venter the plant also develops into a calyptra which protects the sporophyte as it undergoes meitotic division and numerous spores are discharged. After being discharged the spores remove the spore wall and germinate for several days. During germination an asymmetric division occurs forming a smaller cell which produces the first rhizoid. The rhizoid is used for nutrient uptake similar to a root hair cell (Fig.1.7). The larger cells continues to divide multiple times in a single direction to produce the protonema. After development of the protonema the cells then divide irregularly to produce a spherical shape. This spherical shape then produces an apical notch, dividing the thallus, which develops into two nodes and gives the liverwort its characteristic shape (O'Hanlon 1926; Crandall-Stotler 1981; Shimamura 2016). Alongside this, the male and female plants can both reproduce asexually. Multicellular gemmae form in gemma cups on the midrib of the plant from a single apical cell at the base of the cup (Fig.1.7; Barnes and Land (1908)). The dominant haploid life-phase of the plant simplifies the process of genetic manipulation, particularly with respect to generating knock out lines. Also, the plant can be propagated through culturing of gemmae or thallus tissue resulting in a stable genetic line without the need for crossing. The development of

gametangia can be stimulated under far red light also provides a simple a quick way to study the reproductive traits (Miller and Colaiace 1969). Together this makes for an attractive model organism however it was not until recently that the genome was sequenced.

1.4.2 The Genome of *M. polymorpha*

M. polymorpha subsp. ruderalis nuclear and organellar genomes have been sequenced (Birchler-Causse 1993; Bowman 2016; Shimamura 2016) using whole-genome shotgun sequencing. A single female was sequenced from a cross between a wildtype accession Takaragaike-1 (Tak) male line with previously sequenced V chromosome (Yamato et al. 2007) and a Tak-2 female line with an U chromosome introgressed into a Tak-1 autosomal background. Assembly of the contigs of sequence covers 225.8 Mbp of nuclear sequence, 120.3 kbp of plastid sequence and 180.2 kbp of mitochondrial sequence (Bowman et al. 2017). *Marchantia paleacea* has previously published sequences of plastid and mitochondrial DNA and differs from the sequences of *M. polymorpha* (Oda et al. 1992; Ohyama et al. 1986; Kisiel et al. 2011). 19,138 nuclear encoded protein coding genes were predicted to be expressed with 5,385 potential alternative protein coding transcripts also predicted in *M. polymorpha* (Bowman et al. 2017). Single copies were found for a number of regulatory genes and the plant has 9 chromosomes which matches the number of chromosomes found in extant liverwort lineages, which implies the absence of whole genome duplication events (Berrie 1960; Bowman et al. 2017). However, chromosomal analysis of individual genes and gene families is still difficult as the contigs have yet to be mapped to the chromosome (with the exception of the sex chromosomes) and therefore observations on chromosomal location are not possible. So far the V chromosome has 105 annotated genes, mostly involved in sperm motility and development, whereas the U chromosome has 75 annotated genes. 20 of these genes on the U chromosome seem to have homologues on the V chromosome and therefore these genes on different sex chromosome can be considered alleles.

1.4.3 Molecular and Genetic Tools Developed in *M. polymorpha*

M. polymorpha has become an increasingly useful tool for mutagenesis studies with the recent release of its genome (Bowman et al. 2017). A range of genetic tools have been developed for *M. polymorpha* including *Agrobacterium*-mediated transformation and biolistic transformation (Ishizaki et al. 2008; Kubota et al. 2013; Chiyoda et al. 2008; Takenaka et al. 2000). CRISPR-Cas9 in *M. polymorpha* has now been used in several studies to successfully knock out genes (Albert et al. 2018; Flores-Sandoval et al. 2018; Sugano et al. 2018). CRISPR/Cas9 causes targetted double strand breaks and homologous sequences can me made for either side of the break for homologous recombination based gene insertion (Ishizaki et al. 2013a). Both the nuclear and plastid genomes can be modified using the CRISPR-Cas9 technology and therefore a wide range of genetic manipulations can be performed. A range of other techniques have also been developed which are useful for genetic analysis and mutagenesis of the plant, including cryopreservation of spores and gemmae which means lines can be stored for future use (Tanaka et al. 2015). Standard DNA and RNA extraction kits work on *M. polymorpha* and therefore a range molecular analysis can be used including RNA sequencing, qRTPCR and PCR.

1.5 Aims and Objectives of this Project

Ca^{2+} is an important secondary messenger in eukaryotes involved in a wide range of responses to stress as well as governing a number of fundamental growth and developmental processes. In plants, calcium signalling has been implicated in tolerance to drought, salt, heavy metals and biotic stresses such as innate immunity pathways. Calcium signals are also involved in core regulated machinery governing growth, stomatal conduction, nutrient uptake, endosymbiosis with beneficial bacteria, and pollen tube extension in fertilisation. Calcium is bound by a wide range of calcium decoding proteins to carry out function in these processes or to transduce

signals to a number of other signalling pathways as so-called 'master regulators'. A number of these calcium decoding proteins do not have ascribed function due to functional redundancy, whereby more than one calcium decoder overlap in function and single gene knockout cannot be used to determine the calcium decoders role. Numerous groups have tried to propose work-arounds to this issue such as in the case of inducible gene knockout to determine the role of AtCPK10/30/32 function in nitrate assimilation and growth. However, as it is difficult to determine bioinformatically which calcium decoders are functionally redundant new approaches have to be explored. *M. polymorpha* is haploid dominant plant with a published genome and a range of molecular biology tools in active development for its study. Without whole genome duplication and hybridisation events in its history, it is likely to have a reduced calcium decoding toolkit and could potentially have reduced functional redundancy allowing for gene function to be ascribed quickly.

The aim of this project was to investigate the number of calcium decoders of specific families including the CBLs, CIPKs and CDPKs to see if they are reduced in number and hence complexity. If *M. polymorpha* has reduced complexity of calcium decoders in the CBL, CIPK and CDPK families it may be that functional redundancy is less of an issue in this plant than it is in a lot of the angiosperms that are currently studied for calcium decoder function. One aim would be to try and interrogate these protein families in *M. polymorpha* to see if function can be ascribed to a number of these calcium decoders through single gene knockout using CRISPR/Cas9 methods to resolve if this system is free from, or reduced in, functional redundancy. The focus of this functional annotation is on salt stress tolerance which is known as a function for CBLs, CIPKs and CDPKs in plants such as *A. thaliana* and therefore seems to be a core functional purpose for a number of calcium decoder families. Also salt stress tolerance investigation is increasing in economical value as arable land is rising in salinisation due to climate change causing rising sea levels and due to the intensification of farming methods damaging land use. Another aim of this project would be to see if the ascribed function of the calcium

decoder in *M. polymorpha* can be used to inform the original function of calcium decoders through macroevolutionary analysis. Potentially unknown functions of calcium decoders could also be discovered which would provide direction to studies of calcium decoder families in later diverging plants. Alongside this, novel functions may be found for early diverging plant calcium decoders that could be transferred to economically important crop plants for biotechnological purposes.

Simultaneously the work of this project could provide a system for looking deeper at mechanistic function of these calcium decoder pathways. Currently it is difficult to understand how certain calcium decoder such as CBLs and CIPKs function on a mechanistic level as each CBL and CIPK can interact with more than one CBL and CIPK in angiosperms such as *A. thaliana* and the method and requirements for interaction seems to differ between each CBL/CIPK pair. For example SOS3 (AtCBL4) interacts with SOS2 (AtCIPK24) without the need for phosphorylation whereas AtCBL10 is found in different tissue and requires phosphorylation and relocation on a subcellular level to carry out the same function as SOS3 (AtCBL4). Little is known abouth phosphorylation requirements for certain CBL and CIPK interactions as well as how relocations take place. Similarly it is currently not known if CBLs and CIPKs interact at basal calcium levels, or if they require constant interaction to continue CIPK activity after the calcium signal induced activation. The basis of the studies in this thesis could promote *M. polymorpha* as the study system of choice to investigate these details if functional redundancy is found to be reduced.

Chapter 2

Materials and Methods

2.1 Media Used

Table 2.1: Media Formulas used in this project

Media Name	Formula per Litre
LB	LB Broth Miller 25 g (Formedium LMM0105)
LB + Agar	LB Broth Miller with Agar 40 g (Formedium LMM0202)
SOC	SOC Broth 31.5 g (Formedium SOC0201)
YPAD	Yeast Extract 10 g (Formedium YEM02), Peptone 20 g (Formedium PEP01), D(+)-Glucose Anhydrous 20 g (Fischer G/0500/60), Adenine sulphate 20 mg (Formedium DOC0228)
SD-LWHA	1.9 g Yeast Nitrogen Base without Ammonium sulphate and Amino Acids (Merck Y1251), Ammonium Sulphate 5 g (Fischer A/6480/53), D(+)-Glucose Anhydrous 20 g (Fischer G/0500/60), Agar 20 g (Sigma A9539), Yeast Synthetic Dropout -HLWA (Sigma Y2021)
SD-LWH	SD-LWHA, Adenine sulphate 40 mg (Formedium DOC0228)
SD-LW	SD-LWHA, Adenine sulphate 40 mg (Formedium DOC0228), Histidine hydrochloride monohydrate 20mg (Sigma H0755000)
MS + Agar	2.165 g MS with vitamins (Duchefa Biochemie M0222), 8 g Agar (Sigma 05040), 10 g Sucrose (Thermo Scientific J65148.A1)

Table 2.2: Antibiotic Concentrations used in this project

Antibiotic	**Final Concentration (μg/ml)**	**Solvent**
Spectinomycin	200	H_2O
Ampicillin	100	H_2O
Kanamycin	25	H_2O
Gentamycin	40	H_2O
Rifampicin	50	DMSO
Hygromycin	10	H_2O
Cefotaxime	100	H_2O
Carbenicillin	100	H_2O

2.2 Cloning Methodologies and *in vitro* validation

2.2.1 Golden Gate cloning

Cloning of various constructs including for bimolecular fluorescent complementation, yeast two-hybrid, subcellular localisation and CRISPR was carried out using the MoClo system of Golden Gate cloning (GGC; Weber et al., 2011). GGC utilises type IIS restriction enzymes that cleave DNA outside of the recognition site at set distances dependent on the enzyme. Knowledge of the distance of the cleavage site from the recognition site of the restriction enzymes has allowed for the development of standard overhangs that allow for highly efficient cloning and interchangeable parts. Level 0 creates modules with specific overhangs for Promoters, 5' UTRs, Signal peptides, CDSs and Terminators (Hence positions PUSCT; Fig.2.1). The specific overhangs on either side of these derived modules define the position for the next level of cloning. If a 5'UTR is not required promoters can be generated that have the 5' overhang of promoter module and the 3' overhang of a 5'UTR which would then be classified as a 'PU' construct (Fig.2.1) and the same could be true if a signal peptide is not required and the coding sequence of the gene can be cloned as an SC module. Additional overhangs can be generated to to make multipart sections for each module. To clone the signal tag to the N-terminus of the coding sequence an SC1 position was generated with a distinct overhang from the S position

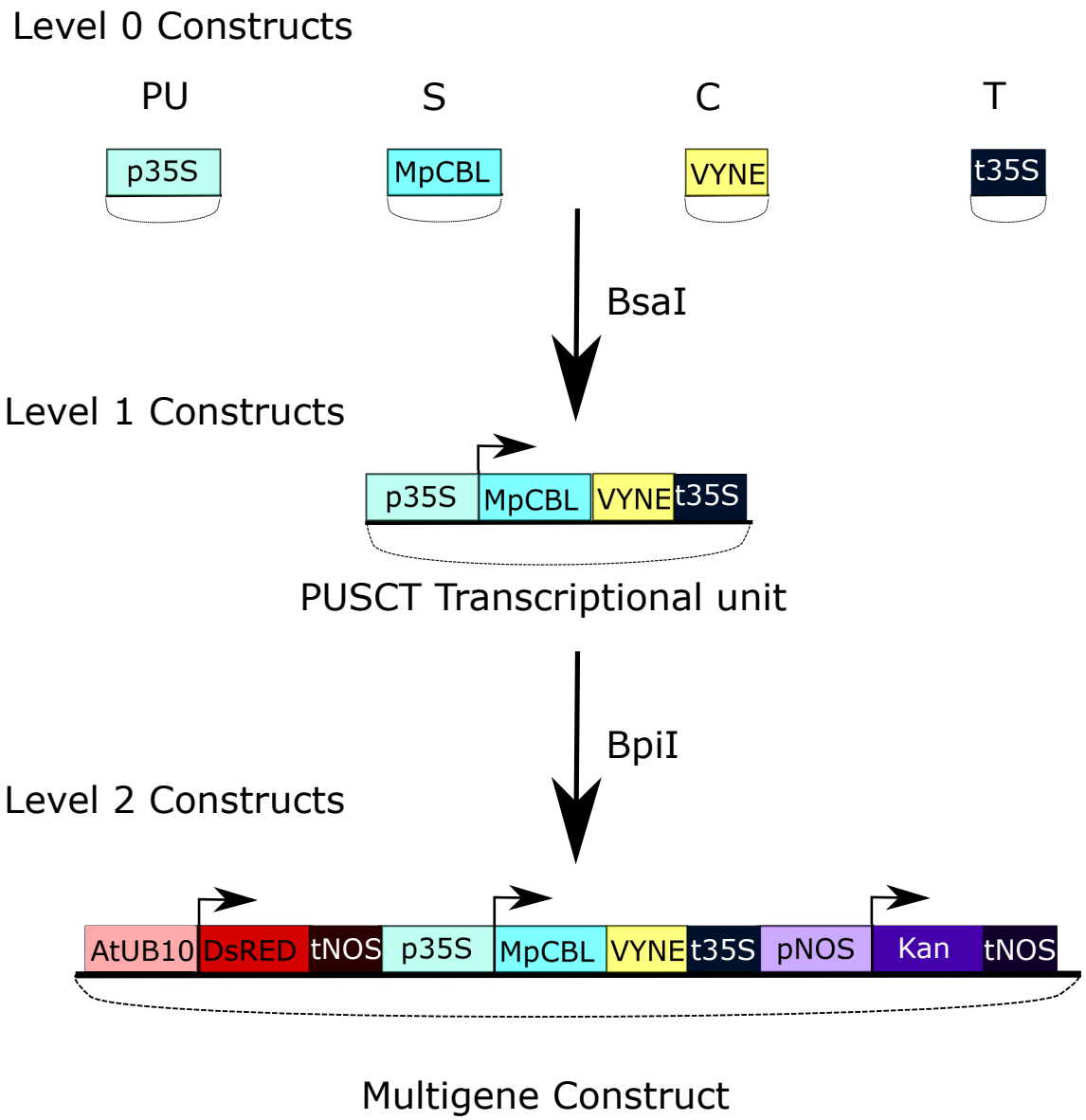


Figure 2.1: Schematics of different levels of Golden Gate Cloning. Level 0 is the different cloning parts in PUSCT positions for Promoter, 5' UTR, Signal peptide, CDS and terminator positions. If not all positions are used positions can be skipped, hence the p35S promoter in PU position. Level 1 constructs are transcriptional units that are placed in plasmids that define directionality and position in the multigene construct. Level 2 constructs are multigene constructs.

C-terminal overhang and the CDS was cloned into the C2 position. Level 1 cloning in this system includes all of the components to make a full transcriptional unit which is then placed into a directional and positional plasmid labelled forward or reverse with positions 1-7, as seven transcriptional units can be included in a single GGC reaction to create a Level 2 construct (Fig.2.1).

Each level of GGC has dual selection with both antibiotic selection and colour

based screening. Level 0 and 1 destination plasmids contain a lacZ cassette that is cleaved out to insert the PCR fragment of choice (Level 0) or transcriptional unit (Level 1) to generate a module and hence blue/white selection can be used to select for white colonies. Level 0 destination plasmids are cleaved with the BpiI type IIS restriction enzyme and allow for antibiotic selection with spectinomycin. Level 1 destination plasmids are cleaved with the BsaI type IIS restriction enzyme and allow for antibiotic selection with ampicillin (Fig.2.1). Level 2 constructs (multigene casettes) allow for antibiotic selection with kanamycin and are cleaved with BpiI type IIS restriction enzyme and the colour selection is generated by a CRed module (for red coloured canthaxanthin biosynthesis) that is cleaved out in cloning hence red/white selection is used (Fig.2.1). Instead of filling all seven positions in the Level 2 constructs 'end-linkers' can be used to seal the construct with fewer parts and these end linkers can include colour casettes with LacZ that can be cleaved out to allow for further cloning additions to the construct, in that case the colour selection would be red/blue and further cloning could be blue/white hence making a level 2-1 construct.

The GGC cloning reaction using the MoClo method is standardised for all levels of construct building with only a difference in restriction enzyme used at each level. The final reaction contains 0.1mg/ml Bovine serum albumin (BSA), 1X T4 DNA ligase buffer (NEB), 1.5 μl T4 Ligase enzyme (NEB), 15 U BpiI or BsaI, 100 ng vector backbone, 100ng DNA per insert (either plasmid or PCR product) and dH20 to make a total volume of 15 μl. This reaction is cycled 25 times between 37°C for digestion by the restriction enzyme for 3 minutes and 16°C for ligation for 4 minutes. As more final product is generated in each cycle and as the final plasmid is digested by a different type IIS restriction enzyme an increasing amount of final plasmid is generated creating a more efficient reaction. The GGC reaction finished with 5 minutes at 50°C and 5 minutes at 80°C to denature the enzymes and 1 μl of final reaction mix can be used to transform bacteria.

2.2.2 PCR for Level 0 Modules

CBLs and CIPKs from *M. polymorpha* were synthesised commercially (Invitrogen GeneArt) in an SC position for level 0 cloning. However mutants with a removal of the NAF domain and remobilisation into different positions for different construct usage was required such as for the ability to add different fluorescent tags or interaction tags for bimolecular fluorescent complementation and yeast two-hybrid analysis. This required the usage of primers containing BpiI resctriction sites and that confer different overhangs so that these CDSs could be cloned into the different plasmids. Polymerase Chain Reaction was used to generate these DNA fragments with Phusion high-fidelity DNA polymerase (Thermo Fisher) with 5'-3' exonuclease activity to reduce the likelihood of introducing SNP mutations. The reaction contained 1X HF Phusion buffer, 0.2 mM dNTPs, 1 μM Forward and reverse primers, 50 ng of template DNA, 0.02U of Phusion High Fidelity DNA polymerase, and water to 20 μl reaction volume. The reactions were heated to 95°C for 30 seconds and then cycled 30 times at 95°C for 10 seconds, 54°C for 20 seconds and 72°C for 1 minute per 1 kb of desired DNA fragment. PCR products were seperated by gel electrophoresis on a 1% agarose gel (Tris-Acetate-EDTA) and visually compared using ethidium bromide post staining (10 mg/ml) compared to a 1KBplus ladder (NEB) to find the correct size fragment. The correct sized fragment was then cut out and extracted using QG buffer (10 minutes at 50°C). 100 μl of isopropanol was added along with 10 μl silica and DNA bound to the suspended silica (2 minutes at room temperature). QG was then removed by centrifuging and pipetting off the waste. DNA bound to silica was washed with column wash buffer (200 μl; Promega) and then eluted by heating to 72C with 20 μl nuclease free water. The DNA in the water was then added as 1 μl addition to a GGC reaction to make level 0 constructs. MpCIPKs with the NAF domain removed were amplified as two parts: A N-terminal part up to the start of the NAF domain and a C terminal part from the end of the NAF domain. These sections were generated with primers (Table.2.8) containing BpiI sites that are cleaved out and the two sections fuse together scarlessly in a GGC reaction to

make a full length CIPK without the NAF domain.

2.2.3 Microbiological strains

Escherichia coli strain DH5α (Invitrogen) was used for cloning of plasmids less than 10 kb. DH10B (Invitrogen) was used for constructs greater than or equal to 10KBp. *Agrobacterium tumefaciens* strain GV3101 (Larebeke et al. 1974) was used for transient expression of *Nicotiana benthamiana* leaves and *Marchantia polymorpha* gemmae. *Saccharomyces cerevisiae* stains Y187 and AH109 were used for yeast two-hybrid construct transformation and subsequent mating to make a diploid strain.

Chemically competent *E. coli* cells

E.coli strains were streaked from glycerol stocks onto an LB-Agar plate and incubated overnight at 37 °C. A colony was selected from the streak plate and innoculated into 5 ml liquid LB overnight at 37 °C. 5 ml of overnight culture was added to 50 ml of liquid LB and incubate at 37 °C until an OD600 of 0.6 was reached then incubated on ice for 15 minutes. Cells were spun down at 4000 rpm for 10 mintues in the cold room and the pellet was resuspended in 20ml $MgCl_2$ (0.1 M). Cells were spund down again (4000 rpm, 10 minutes) and resuspended in 20 ml $CaCl_2$ (0.1M) and and then incubated for 30 minutes on ice. Cells were then pelleted (4000 rpm, 10 minutes) and resuspended in 250 μl $CaCl_2$ (1 M), 1 ml glycerol (50%) and 1250 μl H20 and aliquoted to 100 μl and flash frozen in N_2 for later transformations.

2.2.4 Transformation of *E. coli* competent cells

One microlitre of the GGC reaction was used to transform 20 μl chemically competent *E. coli*. The DH5α or DH10B cells were chilled on ice for 10 minutes with the DNA before heat shock at 42°C for 30 seconds. After heat shock, cells were placed on ice for 1 minute and had 500 μl SOC broth added. Then cells were incubated at 37°C with shaking for 1 hour. The bacteria were then plated on LB-Agar plates containing an appropriate antibiotic (Chapter 2.2.1) and for those that require blue/white

screening X-Gal (40 μg/ml) and isopropyl beta-D-1-thiogalactopyranoside (IPTG; 100μg/ml) was added to provide a substrate and inducer, respectively, for the colour based screen. Plates were then incubated overnight at 37°C and correct coloured colonies that survive antibiotic selection were tested by colony PCR.

2.2.5 Colony PCR

The colony PCR used a backbone primer GG2 for level 0 and level 1 and GG4 for level 2 and a gene specific primer for each construct generated (Table.2.7). The reaction mixture contained: 5 μl of GoTaq G2 Green Master Mix (1X; Promega), 2 μM of forward primer, 2 μM of reverse primer and 3 μl H2O. This reaction mix was then innoculated with a single colony and the reaction was cycled reaction at 98°C for 10 minutes, followed by 30 cycles of 98°C for 10 seconds, 54°C for 20 seconds and 72°C for 1 minute per kb amplified. The PCR reaction finished with a single step of 72°C for 5 minutes. The PCR products are then separated by electrophoresis as previously described (Chapter 2.2.1).

2.2.6 DNA extraction and Glycerol Storage

Colonies with the correct amplified product based on colony PCR were then innoculated in 5 ml LB liquid culture with appropriate antibiotic overnight at 37 °C and 200 rpm. For long term storage 0.5 ml of the overnight culture was added to glycerol to a final concentration of 20% glycerol and stored in -80°C. The remaining culture was centrifuged (4000 rpm for 2 minutes) and supernatant was discarded. The plasmid was then extracted using the Wizard Plus SV Minipreps DNA purification system (Promega) following the manufacturer's instructions. Purified plasmid was eluted in nuclease free water and concentration and purity was measured using Nanodrop 8000 (Thermo scientific). All constructs were then diluted to 100 ng/μl which is then stored at -20°C until use in further tests or GGC reactions.

2.2.7 Restriction Digest Validation

All GGC reaction generated constructs were validated with restriction digestion to verify the size of the insert into the plasmid. To do this purified plasmid DNA was digested for 1 hour at an appropriate temperature for the restriction enzyme. Level 0 constructs were verified with digestion using BsaI to ensure that the liberated fragment was the expected size. Level 1 constructs were digested with BpiI and Level 2 were digested with another appropriate enzyme that can generate 3 fragments such as EcoRI for CRISPR level 2 constructs. Digestion was carried out with 2 μl of appropriate buffer, 10 μl of 100 ng/μl of purified plasmid (hence 1 μg of plasmid), 8 μl of water. Fragment sizes were seperated by gel electrophoresis and assessed to give the expected banding pattern. Cloning was carried out *in silico* using plasmid editing software (ApE) and expected banding patterns were generated using ApE digestion (jorgensen.biology.utah.edu/wayned/ape/).

2.2.8 Sanger Sequencing

To identify correct mutations for delta NAF removal and to ensure that no SNPs were generated during level 0 generation all constructs were validated with Sanger Sequencing by submitting plasmid DNA to Eurofins for analysis. Samples are prepared by adding 800 ng of plasmid DNA, 2 μM of appropriate primer and 8 μl of dH_2O. The output sequence was then compared to predicted sequence using Basic Local Alignment Search Tool (BLAST; Camacho et al. (2009)).

2.2.9 Final Constructs Generated

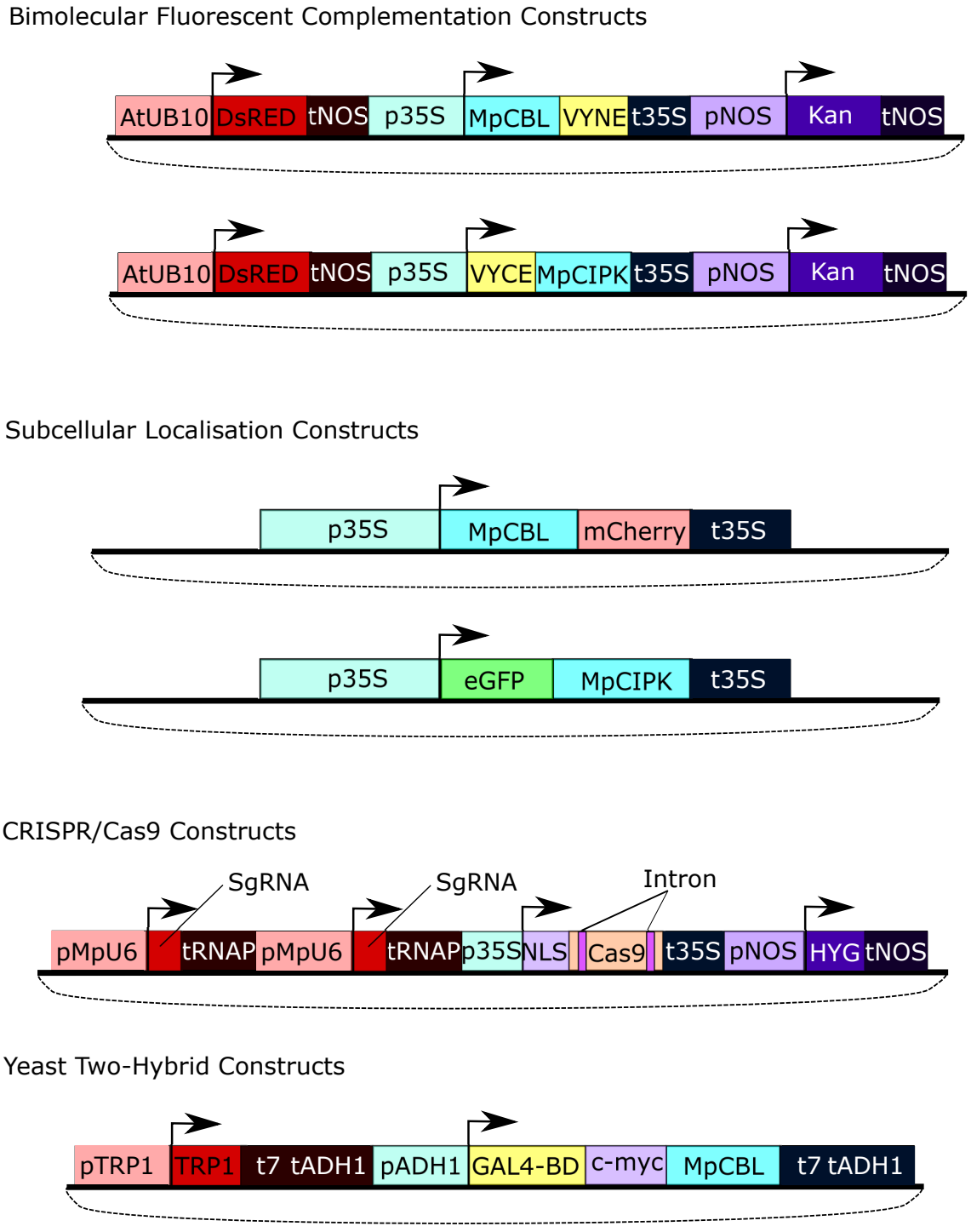


Figure 2.2: Schematics of Final constructs used in this thesis. Top panel shows the two BiFC constructs. Middle panel shows CRISPR/Cas9 construct design. Bottom Panel shows Yeast Two-Hybrid construct design.

2.3 Yeast Two-Hybrid Protocols

The Yeast two-hybrid system relies on a GAL4 transcription factor that can drive the expression of of genes involved in the synthesis of essential amino acids histidine (H) and chemical base adenine (A). The GAL4 transcription factor is split into binding domain, which can bind to the promoter element of these biosynthesis genes, and activation domain, which activates the expression of the genes of interest. Each of these domains are fused to a seperate gene of interest and if the protein products of those genes of interest interact they reconstitute the GAL4 transcription factor which allows yeast cells to grow on the interaction plate lacking these essential compounds.

2.3.1 Yeast Transformation

Yeast two-hybrid constructs generated by GCC with the gene of interest fused to the activation domain are transformed into *S. cerevisiae* AH109 and constructs with the gene of interest fused to the activation domain were transformed into Y187. To prepare cells for transformation the two strains were grown seperately overnight at 30°C in YPAD to make a dense culture. Both *S. cerevisiae* strains were innoculated into fresh 50 ml YPAD and were grown until an OD_{600} of 0.6-0.8. The cultures were then pelleted in a centrifuge (1000 rpm, 5 minutes) and washed with 25 ml of water and pelleted again (1000 rpm, 5 minutes). The pellet of each strain of culture was resuspended in 1.5 ml of lithium acetate (LiOAc, 1 M) to make the prepared cells for transformation. To carry out the transformation a PEG/LiOAc mastermix with DNA was generated with 240 μl 50% PEG (3350), 36 μl LiOAc (1 M), 25 μl salmon sperm single stranded DNA (2 mg/ml) and 3 μl of purified plasmid (100 ng/μl) which was then vortexed for 1 minute. 100 μl of yeast cells of the correct strain was added to the PEG/LiOAc/DNA mastermix and vortexed for 1 minute, then was heated to 30°C for 30 minutes. Cells were then heated 42°C for 30 minutes. Cells were pelleted at 700g for 5 minutes then supernatant was removed and cells were resuspended in 100 μl of water which was then plated on

an appropriate selection plate of SD (SD-L or SD-W). Constructs with the binding domain fused to a gene of interest have a tryptophan synthesis gene in for selection on plates without tryptophan (SD-W, Leucine 100 mg/L) and constructs with the activation domain fused to the gene of interest have a synthesis gene for leucine in which allows for selection on plates without leucine (SD-L, Tryptophan 20 mg/L).

2.3.2 Yeast Mating

Y187 and AH109 strains of *S. cerevisiae* are compatible mating types so Y187 and AH109 harbouring different plasmids can be mated to create diploid yeast with both constructs. As an example the plasmid harbouring MpCIPK-A:BD and a gene for the synthesis of tyryptophan is transformed into Y187, and a plasmid harbouring MpCBL-A:AD and a gene for the synthesis of leucine is transformed into AH109, and the two yeast strains can be mated to generate yeast harbouring both MpCIPK-A:BD and MpCBL-A:AD. As both plasmid have different selection, both plasmids were selected for by removing both leucine and tryptophan from the SD-Agar media (SD-LW). To carry out the yeast mating, the yeast harbouring the plasmids of interest were innoculated in liquid media (YPAD) and the innoculums were mixed and incubated overnight at 30°C with shaking. 10 μl of this mated innoculum was taken and mixed with 90 μl of water which was then plated onto SD-Agar without leucine and tryptophan (SD-LW). These plates were then incubated at 30°C for 3-5 days to select for colonies harbouring both constructs.

2.3.3 Yeast interaction

Each colony resulting from the mating can represent a biological replicate in the yeast interaction test so 5 colonies were selected for 5 biological replicates. The selected colonies were incubated overnight in liquid culture (YPAD) at 30°C with shaking. The overnight culture was then diluted to an OD_{600} of 1.5. The interaction test was carried out on both SD-LW (colony selection plate) and SD-LWH which was the interaction test plate lacking histidine. If the two proteins of interest can

interact and thereby reconstitute the GAL4 transcription factor then the colony will grow without histidine as the transcription factor will drive the expression of the gene that allows for the biosynthesis of histidine. As the strength of the interaction can vary the colonies are plated with 20 μl spots and then diluted 1:10 for 5 more spots and growth of the yeast on each spot was assesed after 4 days when grown at 30°C.

2.3.4 Yeast Pop PCR

All colonies of the yeast two-hybrid were checked for the correct construct using the qRT-PCR primers in a PCR to assess for the CBL and CIPK in the transformed yeast. The colony PCR was carried out by taking a colony of yeast and innoculating into 50 μl water and heating the innoculum to 99 °C for 5 minutes (Horecka and Chu 2017). 1 μl of this boiled innoculum was then used as the basis for a colony PCR and 1 μl less water was used (Chapter 2.2.5).

2.4 Western Blotting

Table 2.3: Western Blot Buffer Formulas used in this project

Buffer Name	**Formula**
SDS-PAGE Running Buffer	For 1 Ltr: Tris 3.28 g, Glycine 14.42 g, SDS 10 ml
Laemelli Buffer	For 20 ml: 0.5M Tris pH 6.8 5ml, SDS 0.8 ml, Glycerol 4 ml, Bromphenol blue 0.004 g, 1M Dithiothreitol (DTT) 4 ml
SDS-PAGE Resolving Gel	For 2 gels: 1.5M Tris pH 8.8 5 ml, dH_2O 7.9 ml, SDS 0.2 ml, Acrylamide/Bis solution 6.7 ml, Tetramethylethylenediamine (TEMED) 10 μl, APS (10 %) 200 μl
SDS-PAGE Stacking Gel	For 2 gels: 0.5M Tris pH 6.8 1.5 ml, dH_2O 3.3 ml, SDS 60 μl, Acrylamide/Bis solution 1.005 ml, Tetramethylethylenediamine (TEMED) 6 μl, APS (10 %) 60 μl
Transfer Buffer	1 Ltr: Tris 4.66 g, Glycine 2.34 g, SDS 1 ml, Ethanol 200 ml, pH 9.2
TBS	1 Ltr: 1M Tris-HCl 10 ml, 5M NaCl 30 ml, pH 7.5
TBS-Tween	1 Ltr: 1M Tris-HCl 10 ml, 5M NaCl 30 ml, Tween20 1ml, pH 7.5
Blocking Buffer	1 Ltr: 1M Tris-HCl 10 ml, 5M NaCl 30 ml, Tween20 1ml (pH 7.5), skimmed milk powder 5 g

Western blotting was carried out on total protein of yeast from liquid culture

that was used in the interaction tests. Each of these yeast strains harbour the two plasmids required for each interaction. As a control, yeast harbouring a single construct for CBL-A:AD, and yeast harbouring a single construct for CIPK-A:BD were also used, alongside the yeast strains AH109 and Y187. To ensure that protein was expressed, the western blot was used to detect protein product of each CBL and CIPK in each interaction test. Detection of each protein ensures that none of the negative interactions were due to errant protein degradation or lack of expression.

2.4.1 Protein extraction

All 5 biological replicates used for each interaction test were grown at 30 °C in a 5 ml selective culture (SD-LW and carbenicillin 100 μg/ml) overnight to make a dense culture. This selective 5 ml culture was added to a 45 ml culture of YPAD (and carbenicillin) and grown for 2-6 hours for an OD600 of 0.4-0.6 and then pelleted by centrifugation (10 minutes, 1000 rpm). Pelleted yeast were then washed by resuspension in 20 ml of water followed by pelleting (10 minutes, 1000 rpm). OD units were calculated by multiplying the final OD_{600} by the ml of culture. Per OD unit, 10 μl of laemelli buffer was added to the pelleted cells to make the same amount of yeast in laemelli buffer for each interaction. Yeast and laemelli buffer were then boiled at 95 °C for 30 minutes to to lyse the cells so that protein could be assessed.

2.4.2 SDS PAGE

SDS-PAGE gel were made as described (Table.2.3) with 1.5 mm thickness. A resolving gel of 10 % acrylamide solidified with an isopropanol cover for 45 minutes and stacking gel of 5% acrylamide with a 15 well comb was allowed to solidify for 1 hour. Plates were added to the gel tank after solidification, and running buffer was added to cover the gels. Protein ladder was added to the first well (5 μl) followed by 20 μl of sample in each subsequent well. Each gel had controls of WT (AH109) yeast and yeast containing single constructs for CBL-A, and CIPK-A, which act as positive or negative controls for immunoblot detection later. The SDS-PAGE was

run at 130V for 2 hours to allow for sufficient resolution of proteins.

2.4.3 Wet Transfer to Polyvinylidene difluoride (PDVF)

Wet transfer was used to transfer proteins from the SDS-PAGE gel to activated PDVF membrane. PDVF was activated by incubating in 99 % ethanol for 15 minutes, then washing with dH_2O for 5 minutes and incubated in transfer buffer for 15 minutes. The stacking gel was removed and the SDS-PAGE was placed on two Whatman papers. The activated PDVF was placed on the SDS-PAGE gel followed by two more Whatman papers. This stack was clamped and placed in the transfer tank which was filled with Transfer buffer (Table.2.3). The transfer was performed at 100V for 1 hour and 20 minutes in 4 °C.

2.4.4 Blocking and Antibody Use

Membranes were blocked with TBS-T with milk (5%) for 1 hour post transfer and then washed with TBS-T. Membranes were then incubated overnight in the cold room with the 10 ml of TBS-T with either anti-c-myc-peroxidase (1:10,000; Sigma A5598), or anti-HA (1:10,000; Sigma H6908). Anti-c-myc-peroxidase contained an already conjugated peroxidase so did not require incubation with a secondary antibody. Anti-HA gels were washed 3 times with 20 ml TBS-T then incubated with the secondary antibody anti goat-peroxidase produced in rabbit (1:10,000; Sigma A0545) in TBS-T for 1 hour at room temperature. Peroxidase activity was measured for detection of protein using ECL Select Western Blotting Detection Reagent (Cytiva Amersham) following the manufacturer's instructions.

2.5 Plant Phenotyping and Transformation

2.5.1 Plant Materials and Accessions

Four accessions of *M. polymorpha* were used in this study including Takaragaike 1 (Male), 2 (Female), Cambridge 1 (Male), 2 (Female) supplied by the Haseloff Labra-

tory (Cambridge). All four accessions of *M. polymorpha* were used for phenotyping and the Cam2 accession was used for transformation with CRISPR/Cas9 constructs for genetic knockout. *Nicotiana benthamiana* was used to carry out transient transformation for bimolecular fluorescence complementation studies of MpCBL and MpCIPK proteins.

2.5.2 Plant Growth and Phenotyping Conditions

M. polymorpha were grown through vegetative propogation on 1/2 MS + sucrose (1%) + agar (0.8%) plates (pH 6.0) at 23 °C with a 16 hour photoperiod. Salt condition phenotyping was initiallty carried out for 10 days at 0, 50, 100, 150, 200 mM NaCl for initial test using 10 thallus cuttings. For a more consistent growth gemmae were tested by growing for 7 days on 1/2 MS + sucrose (1%) + agar (0.8%) followed by 7 days on the phenotyoping plates including 0, 50, 150, 200 mM on NaCl or KCl and 100, 200, 300, 400 mM sorbitol for isoosmotic effect. The final phenotyping conditions tested were for thallus cut to 5x5 mm lobes which were plated on 0, 50, 100, 150 mM NaCl as this demonstrated the most consistent growth based response. All phenotyping was conducted on MS plates above at 23 °C with a 16 hour photoperiod. $ZnSO_4$ was used to phenotype Cam2 with concentrations of 0.1, 0.5, 1 and 5 mM using thallus cuttings on 1/2 MS + sucrose (1%) + agar (0.8%) for 1 week.

Table 2.4: Supplementation of NaCl to MS plates to attain phenotyping concentrations

Final [NaCl] mM	**Water (μl)**	**NaCl added (μl; 5M stock)**
0	2000	0
50	1800	200
100	1600	400
150	1400	600
200	1200	800

Table 2.5: Supplementation of ZnSO4 to MS plates to attain phenotyping concentrations

Final [ZnSO4] mM	**Water (μl)**	**ZnSO4 added (μl; 50 mM stock)**
0	2000	0
0.1	1960	40
0.5	1800	200
1	1600	400
5	0	2000

2.5.3 Preparing chemically competent *Agrobacterium tumefaciens*

GV3101 strain of *A. tumefaciens* was streaked on LB-Agar with rifampicin and gentomycin from glycerol (Table.2.2). A single colony was selected and grown in 5 ml liquid culture at 28 °C in LB with rifampicin and gentamycin (Table.2.2). 0.5 ml of the 5 ml culture was transferred into a 50 ml LB liquid culture and incubated at 28 °C with shaking until an OD_{600} of 0.6-0.8 is reached. Cells were then placed on ice for 10 minutes and were then pelleted using a centrifuge at 4 °C (3000rpm, 10 minutes). Cells were then resuspended in 1 ml $CaCl_2$ (20 μM) and a final concentration of 20% glycerol and were flash frozen in N_2 and aliquots were placed in the -80 °C for later use.

2.5.4 Freeze-Thaw Transformation of *Agrobacterium tumefaciens*

GV3101 *A. tumefaciens* was transformed by adding 300 ng of plasmid DNA to 25 μl of thawed chemically competent cells and mixed by tapping. Cells were frozen in liquid nitrogen and thawed at 37 °C for 5 minutes. SOC broth (500 μl) was added to the cells as a nutrient media to allow for growth. The culture was then inbucated at 28 °C with shaking (300 rpm) for two hours. The transformed *A. tumefaciens* was then plated onto LB-agar plates with appropriate antibiotic selection for the strain (rifampicin and gentamycin; Table.2.2) and for the plasmid (kanamycin; Table.2.2). Plates were incubated at 28 °C for 3 days. Positive colonies were selected using

colony PCR (Chapter 2.2.5).

2.5.5 Transient Transformation of *N. benthamiana*

N. benthamiana leaves were transformed with the *A. tumefaciens* strain GV3101 harbouring bimolecular fluorescence complementation (BiFC) constructs to test for protein interactions. *N. benthamiana* plants were grown for 4 -weeks in a controlled environment at 23 °C with a 16 hour photoperiod. 3 colonies of PCR positive *A. tumefaciens* were selected and innoculated together in a 10 ml culture of liquid LB with rifampicin, gentamycin and kanamycin (Table.2.2) for BiFC constructs. GV3101 harbouring BiFC constructs were incubated at 28 °C with shaking (180 rpm). The *Agrobacterium* culture was then centrifuged at 3000 rpm for 10 minutes and resuspended to an OD_{600} of 0.5 in $MgCl_2$-MES (10 μM concentration for both) with 200 μM of acetosyringone. The culture was left for 2 hours at room temperature in the dark and then infiltrated with a needleless syringe in 3 leaves for 3 technical replicates. After 72 hours the transformed material was harvested using a leaf punch and fluorescence was assessed on a confocal microscope (Zeiss LSM980-Airyscan). Leaf discs were placed on slides to inspect the underside of the leaf at 20X magnification with a 561 nm laser for excitation at 20 % power.

2.5.6 Stable *Agrobacterium*-Transformation of *M. polymorpha*

CRISPR/Cas9 constructs were transformed into *M. polymorpha* using *A. tumefaciens* GV3101 suspended in $MgCl_2$-MES (10 μM) with 200 μM of acetosyringone left in the dark at room temperature for 4 hours. Gemmae of *M. polymorpha* were grown upside down on MS-Agar media (1/2 MS; 1% sucrose; 0.8 % Agar) for 3 days and then placed in 6-well plates and sealed with parafilm sealing with the *Agrobacterium* co-culture ($MgCl_2$-MES + acetosyringone) at 23 °C in the dark for 3 days with shaking at 120 rpm. Plantlets were then washed 3 times with water containing cefotaxime (Table.2.2). Washed plantlets were then plated onto MS-Agar plates

(1/2 MS; 1% sucrose; 0.8 % Agar) with hygromycin and cefotaxime (Table.2.2). Hygromycin to select for positive transformants and cefotaxime to remove *Agrobacterium*.

2.5.7 Plant DNA and RNA extraction

Plant DNA was extracted for verifying the genotype of plants transformed with CRISPR/Cas9 constructs and grown on hygromycin selection using a DNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. The resulting extracted DNA was assesed by Phusion High-Fidelity PCR of the gene of interest (CIPK-A: c245/c232; CIPK-B: c60/c227). The PCR product was confirmed with Sanger sequencing using 1 μl of DNA, 2 μM of primer (Table.2.7) and dH_2O using forward and then reverse primers. RNA was extracted using RNeasy Micro Kit (Qiagen) from phenotyped plants to assess for changes in expression in genes of interest in Tak1 and Cam2 plants. DNase treatment was used to remove DNA with AMbion Turbo DNase treatment following manufacturer's instructions. RNA quality was assessed by running on a gel and inspecting the ratio of ribosomal RNA. PCR was carried out to ensure no DNA could be amplified after DNase treatment (Chapter 2.2.5; Table.2.6). Quantification of RNA was carried out using a Nanodrop 8000. cDNA was then synthesised using oligo-dT17 which anneals to the polyA tail of mRNA, followed by reverse transcription with Superscript II. RNases were inhibited with RNAsin according to manufacturer's instructions.

2.6 qRT-PCR for expression of Calcium decoders

Expression of genes of interest was measured using the delta delta Ct method on an AriaMx Real-time PCR system (Agilent) using Sybr Green JumpStart Taq ReadyMix (Sigma-Aldrich). The reaction mixture consisted of a 10 μl reaction volume with 5 μl Sybr Green, 2.6 μl MgCl2 (25 mM), 0.2 μl (10 nM) gene-specific forward primer, 0.2 μl (10nM) gene specific reverse primer and 2 μl cDNA template

(1:15 v/v dilution). Reaction conditions were as follows: 95 °C for 4 minutes then 40 cycles of 30 seconds at 94 °C, 30 seconds at 55 °C and 30 seconds at 72 °C. Fluorescent reading was taken after the 55 °C step.

Quality control was carried out on all the primers used in the study. RT-PCR was performed using colony PCR conditions (Chapter 2.2.5) with cDNA as a template. PCR products were assessed through electrophoresis on a 2 % agarose gel to check for a single band of the expected size. As multiple bands could be the same size, melt curves for amplification were also checked on the qPCR machine for a single melt peak which implies a single product as a result of the reaction conditions. Primers were also assessed for efficiency from 10^{-1}-10^{-8} ng/μl template DNA and primers with less than 95% or more than 105% efficiency were removed. Efficiencies were included in the delta delta Ct calculations and used to assess expression changes. All genes and conditions assessed used averaged Ct for three technical replicates with three biological replicates and normalised to two different housekeepers including adenine phosphoribosyl transferase and actin from the same biological replicate with three technical replicates.

2.7 Bioinformatics and Statistics

2.7.1 RNA Sequencing

Data for *M. polymorpha* treated with 100 mM NaCl for 0, 24, 48 hours RNA sequencing were downloaded from the Sequence Read Archive (NCBI; SRA study: SRP268513) with 3 biological replicates using paired end reads (Wu et al. 2021). Data for *M. polymorpha* 50 mM salt treatment for 7 days were taken from Sequence Read Archive (DDBJ; SRA: DRA006681) with 3 biological replicates and paired end reads (Tanaka et al. 2018).

Prior to quantification an index file was made for selective alignments. Genome targets were selected and a decoy file was generated to detect spurious alignments (Grep) and the entire genome (v5.1, marchantia.info) was concatenated to the tran-

scriptome for indexing with Salmon (v0.14.1). Paired end reads were quantified by mapping to the Salmon index with Salmon quant utilising validateMappings which validates mapping using a SIMD-parallelized Ksw2 library using a Linux Bash environment. Tximport was used to import quantification in R and a TxDB was used to generate gene IDs from a GFF file (v5.1, marchantia.info). A DESeqDataSet object was generated and DESeq2 was used to check for statistically significant differences between treated and untreated conditions.

2.7.2 Family Finder Script

FamilyFinder is a script that was generated to identify members of a protein family in a protein database using the input of the same protein family from *Arabidopsis*. Edel and Kudla (2015) used BLASTp of *A. thaliana* CBLs against the *M. polymorpha* version 3.1 genome followed multiple sequence alignment (MSA) and manual pruning based on phylogeny and protein domain presence (Pfam). However to carry out such a analysis and expect to find all the members of a protein family, a diverse mix of CBLs would be required for the BLAST input which would also increase the output sequences that need to be filtered. Alongside this, the process would be time consuming to manually analyse in the case of wanting to understand the diversity of multiple protein families. FamilyFinder is a script that automates the BLAST, MSA and protein domain scan process and also carries out some of the filtering to remove repeat sequences and analysis of High-scoring segment pairs.

FamilyFinder script runs as an executable with three arguments (Appendix.A.1). Argument 1 defines the input file which is held in the input folder. The second argument defines a tag for the multiple output files which could be the protein family name such as 'CIPK' to receive CIPK.txt, CIPK.txt.tree and CIPKalign.txt as output files. CIPK.txt includes the filtered sequences from the BLAST alignment. CIPK.txt.tree holds a tree file based on the multiple sequence alignment. CIPKalign.txt contains the trimmed sequences from the multiple sequence alignment.

The first part of the FamilyFinder script focusses on using local BLASTp against

a protein database. The input must be in fasta format and can be one or more sequences. For protein family finding in *M. polymorpha* the input used was the protein sequences of the respective family in *Arabidopsis*. The blastp command runs a local BLAST with an e-value cut-off of 1e-50 using the input against a local BLAST database such as the *M. polymorpha* predicted protein sequences (Bowman et al. 2017) formatted with makeblastdb (Camacho et al. 2009). The output is exported as tab delineated with sequence IDs and sequences. Sed is used to add ' ' to the sequence IDs so that they can be read in fasta format and non-unique sequences are removed. Gaps are then removed from the alignments so that the amino acid sequence of each output can be filtered further (Appendix.A.2).

One aspect by which the *M. polymorpha* potential calcium decoder sequences can be filtered is by length. Many of the calcium decoders have set domains such as four EF hands and set distances between them. The conservation of size of the whole protein sequence means that it is a good aspect by which to filter as many sequences in the output will be too small to contain the correct domains at the defined distance. A python subscript is used to read the input file and determine the minimum and maximum length of any sequence in that file. The maximum and minimum have a buffer added of 20 amino acids to allow for some variability in length in case of an input file with few sequences (Appendix.A.3). The sequences that fit in that range from the BLAST output are then output into another file for the next part of the bash script.

Multiple steps were involved in the multiple sequence alignment by Mafft. Firstly, multiple sequence alignment was carried out on just the filtered BLAST output. This output was then trimmed to allow for a tighter alignment of similar sequences using trimal automated program (Capella-Gutiérrez et al. 2009). The trimmed output was then reformatted for multiple sequence alignment using a bash script modified from (Shaiber, A. 2019). This bash script adds '@' at the end of every ID, then replaces ' ' with '#'. tr -d allows the deletion of a character in this case the newlines followed by another tr command which trades characters '#' for a newline and '@'

for a tab. Sort -unique removes the duplicated sequences in the file. After this the ' ' is added back to label the IDs and the tabs are exchanged for newline which is then written into a new file in fasta format. This file is concatenated with the input file and another comparison file input which was set as the Calmodulins from *A. thaliana* for the calcium decoders but this could be added as a fourth argument to scan other protein families. After this another mafft multiple alignment is used with the trimmed sequences to have an accurate phylogeny based on the areas that match (Appendix.A.4).

2.8 Primers used in the study

2.8.1 qRT-PCR Primers

Primers used for qRT-PCR Analysis were designed using Primer BLAST against the transcriptome of *Marchantia polymorpha* (v5.1, marchantia.info). All primers amplify less than 200 bp and uniquely target the gene. Primers for ADP housekeeper were taken from Saint-Marcoux et al. (2015) and primers for ATP housekeeper were taken from Sharma et al. (2014).

Table 2.6: Primers used for qRT-PCR of the CDPKs, CBLs and CIPKs

Primer Target	Primer ID	Primer Sequence
CBL-A Forward	C 108	AGCGGAAAGAGGTGAAACGG
CBL-A Reverse	C 109	GAGAGGGATGCTGCTGAACC
CBL-B Forward	C 112	GGGCTGCTTCAGCTCAAAAC
CBL-B Reverse	C 113	CGCAAGCTGGAACTCTTCCT
CBL-C Forward	C 116	CAAGTGCTCCACCAGAGGAC
CBL-C Reverse	C 117	GCCTCCGCAAATGTCTTGTC
CIPK-A Forward	C 120	AAACACCCTGCGAACGAGAT
CIPK-A Reverse	C 121	ACCTCAAACACCTCTGTGGC
CIPK-B Forward	C 124	CCTGTACGGATGCACGATGA
CIPK-B Reverse	C 125	AGAACGGAAAGGTTGAGCCC
CDPK-A Forward	C 71	GCGGTCTGAAAAGAGTGGGA
CDPK-A Reverse	C 72	CGAAGTAGGAGAAAGCCGCA
CDPK-B Forward	C 75	AAAAAGATGGCATTGCGGGT
CDPK-B Reverse	C 76	GACATCAGCAGCTTCCATTAGC
CDPK-C Forward	C 79	AACAGAACCCAAAGGAGCGAT
CDPK-C Reverse	C 80	CGGCGATAACCTTCAGAGCA
CDPK-E Forward	C 89	GGGCGAGACTTACTGCATCA
CDPK-E Reverse	C 90	TCTTCTCCGTCCAGAGTGCT
CDPK-F Forward	C 166	CCGGAAGTGGATGTCTGGAG
CDPK-F Reverse	C 167	TGACGTACCAGTGACTTGGC
CDPK-G Forward	C 168	CCCGAAGCCGATGTATGGAG
CDPK-G Reverse	C 169	CCGGCAGCGAGATAGACTTC
LEA1 Forward	BM234	GCTAACAGACCCAGGTGAC
LEA1 Reverse	BM235	TGTTTCCAACGGCAGAGTG
Actin Forward	BM171	GAGCGCGGTTACTCTTTCAC
Actin Reverse	BM172	GACCGTCAGGAAGCTCGTAG
ADP Forward	C 170	CGAAAGCCCAAGAAGCTACC
ADP Reverse	C 171	GTACCCCCGGTTGCAATAAG

2.8.2 Genotyping Primers

Primers used for genotyping of CRISPR/Cas9 knockout lines for CIPK-A and CIPK-B in *M. polymorpha*. Primers were designed for >50bp upstream and downstream of the target sequence of the sgRNAs so that exon 1 could be amplified. The forward and reverse primers were then used to sequence the PCR product.

Table 2.7: Primers for Genotyping CIPK Mutants

Primer Target	**Primer ID**	**Primer Sequence**
CIPK-A Genotype Forward	C 245	GAACATTATGGTTGAC
CIPK-A Genotype Reverse	C 232	CAACCGAAGTTAGTG
CIPK-B Genotype Forward	C 60	ATGGTGGTCAGAAAGG
CIPK-B Genotype Reverse	C 227	GACAGGCAAGGAAAC
GoldenGateL0-1	GG 2	CCTATAAAAATAGGCGTATCACG
GoldenGateL2	GG 4	GCGGACGTTTTTAATGTACTG

2.8.3 Cloning Primers

Primers were used to clone the CIPKs without the NAF domain. Gene specific forward primer (CIPK-A: c178; CIPK-B:c179) were used with a gene specific primer that anneals before the NAF domain (CIPK-A: c174; CIPK-B: c179) for part 1 of the construct. Part 2 of each construct used a gene specific primer for after the NAF domain (CIPK-A: c175; CIPK-B: c177) (Primers: Table.2.8). The two amplicons were joined by GGC (Chapter 2.2.1) using BpiI and into a SC-position plasmid.

Table 2.8: Primers for Cloning CIPKΔNAF from synthesed DNA

Primer Target	**Primer ID**	**Primer Sequence**
CIPK-A Forward	C 178	tgtgaagaccaAATGAGTTTCGCGACTAC
CIPK-A Reverse No NAF	C 174	tgtgaagaccaTCCTCCTTCTTTTCTTG
CIPK-A Forward No NAF	C 175	tgtgaagaccaAGGAAGAACAACCAG
CIPK-A Reverse	BM287	CCCGAAGACTCaagcTCATTTACTG
CIPK-B Forward	C 179	tgtgaagaccaAATGGTGGTCAGAAAG
CIPK-B Reverse No NAF	C 176	tgtgaagaccaTCTTTAGATTCTTTCTGC
CIPK-B Forward No NAF	C 177	tgtgaagaccaAAGATAGACGTCAGG
CIPK-B Reverse	BM289	CCCGAAGACTCaagcTTATTGGCTC

Chapter 3

Discovery of *M. polymorpha* Calcium Decoder Genes

3.1 There are Calcium Decoding proteins in *M. polymorpha*

It has been hypothesised that the calcium signalling toolkit expanded with the emergence of plants onto land, alongside the diversification of other signalling processes. Calcium signalling protein components have been identified in chlorophytes, charophytes and a range of early-diverging land plants (Edel and Kudla 2015). Therefore, the Ca^{2+} signalling toolkit did exist prior to land expansion. The number of representatives of each protein family has expanded along with the diversity of calcium decoding protein families, although this expansion has not been uniform across green lineages. Of the chlorophyta, *Chlamydomonas reinhardtii* had 15 CDPKs but no CCaMKs, CIPKs or CBLs, and *Ostrococcus tauri* only had 3 CDPKs but had both a CIPK and CBL. Whereas the charophyta, *Klebsormidium flaccidum*, had 12 CDPKs, 1 CIPK and 3 CBLs representing a higher diversity in the Ca^{2+} toolkit. The bryophyte *Physcomitrella patens* could have been a good option for understanding calcium decoding but with 25 CDPKs, 2 CCaMKs, 7 CIPKs and 5 CBLs, it may have the same issue of functional redundancy seen in higher plants.

Studies on the CBLs in *P. patens* have already demonstrated that a number of the CBLs are cognate pairs which are likely functionally redundant probably caused by a genome duplication event (Kleist et al. 2014). Therefore the two genome duplication events *P. patens* has gone through likely led to both rapid expansion and a lot of functional redundancy in CBLs but it still could be a good system to understand the mechanisms that cause functional redundancy and likely neofunctionalisation of calcium decoders in angiosperms. However other mosses, hornworts and liverworts may be better models to understand core mechanisms and early functions of calcium decoders without the hinderance of functional redundancy, provided there is no evidence of genome duplication.

Marchantia polymorpha is a early-diverging land plant that is not thought to have undergone any genome duplication events and has its genome published (Bowman et al. 2017). The published genome, along with the rapid development of genetic tools could allow *M. polymorpha* to be leveraged for greater understanding in a range of signalling pathways. For *M. polymorpha* to be useful in understanding calcium decoding it would require representatives in different calcium decoding families with few members in each of the subgroups of the families. Edel and Kudla (2015) interrogated the genome and identified 3 CBLs and 2 CIPKs which differs from *Arabidopsis* and other higher plants that have more CIPK representatives than CBLs. Of the MpCBLs described, two are predicted to localise to the plasma membrane, and one to the tonoplast, which represents the two largest subgroups in higher plants. The MpCIPKs described did not include representatives from the intronless clade as the intronless clade seemingly arose and went through a rapid expansion in the seed plants (Zhu et al. 2016).

The number of MpCDPKs in the genome has not been resolved yet but some can be located in the literature. Bowman et al. (2017) ontology labelled a MpCDPK located on the sex chromosomes when the genome for *M.polymorpha* was sequenced. Nishiyama et al. (1999) described two spliceoforms of a MpCDPK in the sexual structures of the plant with differing exon 6. Sharma et al. (2013) tested another

MpCDPK as a potential qRT-PCR reference gene, but it demonstrated variable expression through the lifecycle of *M. polymorpha*. None of the previously described CDPKs have the same sequence and therefore 3 CBLs, 2 CIPKs and 3 CDPKs were described in literature for *M. polymorpha*. However there is not reason to believe that this represents the complete number of MpCDPKs as no specific identification of MpCDPKs has been carried out and more candidates may be identified in version 5.1 of the genome. Hence, a more thorough analysis of the genome for calcium decoding proteins is required to check how many other representatives may present.

3.1.1 EF hands may reveal the function of calcium decoders

EF hands are one of the most conserved domains found in 71 non-redundant protein families (Sillitoe et al. 2013) and are observed through all domains of life (Nakayama et al. 2000). Theoretically EF hands with more similar calcium binding properties would have more conserved protein identity and therefore be more likely to respond to similar calcium signals. Hence, EF hands will be compared between the discovered calcium decoders in *M. polymorpha*. As the effect of individual residue substitutions is unknown on calcium binding properties, amino acid sequence similarity will not used for the comparison. Comparison of EF hands may reveal more than just calcium binding properties as, while EF hands have a high specificity for calcium, some EF hands can bind other physiologically relevant divalent cations such as Mg^{2+} and Mn^{2+} (Silva et al. 1995; Sánchez-Barrena et al. 2005). However upon binding of other divalent cations the EF hand is not thought to change conformation and therefore these other cations may simply modulate calcium signals through competitive inhibition such as in myosin (Silva et al. 1995). SOS3 (AtCBL4) seems to bind Mn^{2+} with the fourth EF hand which may be for delivery to the CIPK as a preferred cofactor for kinase activity (Sánchez-Barrena et al. 2005). When comparing EF hands, the loop sequences are the most diverse area and directly bind calcium but they are not the only area important for EF hand function.

The helices of the EF hands are also important for determining function as it is

the movement of the helices in relation to each other which results in conformational change in response to calcium binding. The helices also extend the loop sequence from 12 to over 29 amino acids (aa) and creates area of homology either side of the loop which will be useful for domain identity and similarity scoring. If the EF hands are conserved they may be usable to define the highest shared identity with EF hands from a calcium decoder in *A. thaliana* to a calcium decoder in *M. polymorpha*. This can be used to confirm the subgroup based on shared identity of EF hands and also provide a single calcium decoder that shares highest identity which can be used to propose the function of the *M. polymorpha* calcium decoder.

3.1.2 Bioinformatic Analysis of Calcium Decoders from *M. polymorpha* to investigate potential use as a Model System to Understand Mechanisms of Calcium Decoding

Investigating the gene numbers for each of the CBLs, CIPKs and CDPKs in *M. polymorpha* can give an indication of functional redundancy as generally fewer genes in the calcium decoding toolkit indicates less redundancy. However, this is not always true, as has been shown in *P. patens*, as many of the CBLs form cognate pairs. Some of the cognate pairs demonstrated traits that indicate the pair-mate has become a pseudogene such as lack of detectable expression such as PpCBL6 (Kleist et al. 2014). However, in some cases it is likely the pair-mate is functionally redundant. To truly investigate the likelihood of reduced functional redundancy both number of calcium decoders and diversity of the identified decoders has to be assessed.

Diversity can be assessed bioinformatically in a number of ways including looking at phylogeny, protein sequence identity and protein sequence similarity. High levels of identity and similarity may indicate a likelihood of functional redundancy. Phylogeny can be a useful tool to demonstrate which calcium decoder genes in *A.*

thaliana are similar to those found in *M. polymorpha* and therefore have a likely similar function. As calcium decoders fall into subgroups based on function, and localisation in the case of CBLs, it can also be used as a potential indicator of the range of functions for the calcium decoders, in *M. polymorpha*. Alongside this, details such as length, identifier and intron number will be investigated to provide some potential evolutionary context of each calcium decoder, but until the *M.polymorpha* genes are physically mapped to chromosomes this analysis will be limited. Instead intron number can be used as a confirmation of calcium decoder subgroup. The EF hand analysis will be implemented to try and find individual calcium decoders in *A. thaliana* that are 'most similar' and later it will be compared to functions ascribed to knockouts of genes to see if this analysis could be useful in the calcium decoding field to imply function from phylogeny.

3.2 Results

3.2.1 There are 7 CDPKs in *M. polymorpha*

Table 3.1: Characteristics of CDPKs from *M, polymorpha* includings gene IDs for different versions of the genome (V3.1 or V5.1) and length.

CDPK	**V3.1**	**V5.1**	**Chr**	**Length (aa)**	**Length (bp)**
CDPK-A	Mapoly0214s0015	Mp4g13490		476	1431
CDPK-B	Mapoly0007s0233	Mp3g02440		654	1965
CDPK-C	Mapoly0089s0009	Mp3g22080		548	1647
CDPK-C2	Mapoly0089s0009	Mp3g22080		549	1650
CDPK-D	Mapoly0017s0021		U	527	1584
CDPK-E	Mapoly0098s0040	Mp4g01600		565	1698
CDPK-F	MapolyYA0006	MpVg01160	V	527	1584
CDPK-G	Mapoly0004s0215	Mp3g14560		486	1461

Using the FamilyFinder script, 9 gene IDs of potential CDPKs were identified in *M. polymorpha* genome version 5.1 (Table.3.1; Marchantia.info). The sequences retrieved from the script were then analysed with PrositeScan (De Castro et al. 2006) and two had fewer than 4 EF hands and were therefore classified as CDPK-Related Kinases; leaving 7 CDPKs in *M. polymorpha*. As per the *M. polymorpha*

nomenclature these CDPKs were labelled as MpCDPKs (Bowman et al. (2016); Fig.3.1) and until comparisons can be drawn to existing CDPKs they have been labelled as A-G instead of with numbers. MpCDPK-A/B/G are in subgroup I, MpCDPK-C is in subgroup II, MpCDPK-D/F are in subgroup III and MpCDPK-E is in subgroup IV. Of these, MpCDPK-D/F are located on the sex chromosomes and as *M. polymorpha* is haploid dominant; Tak1 (males) have MpCDPK-F, and Tak2 (females) have MpCDPK-D respectively. Therefore *M. polymorpha* has single CDPK representatives in all subgroups with three respresentatives in subgroup I. Typically functionally redundant CDPKs are in the same subgroup and therefore all of the subgroups with single representatives are unlikely to be functionally redundant. However CDPK-A/B/G may have some functional redundancy as they are all members of the same subgroup. MpCDPK-D/F may be functionally redundant but do not coexist in the same adult plant due to its haploid dominant life phase. Therefore there is a wide diversity of CDPKs in *M. polymorpha* with little likely functional redundancy to hamper research into function except potentially subgroup I containing MpCDPK-A/B/G.

Table 3.2: The protein sequence identity of the 7 putative CDPKs from *M. polymorpha* from pairwise alignment using EMBOSS Water.

	CDPK-A	**CDPK-B**	**CDPK-C**	**CDPK-D**	**CDPK-E**	**CDPK-F**	**CDPK-G**
CDPK-A	100						
CDPK-B	76	100					
CDPK-C	67	58	100				
CDPK-D	56	59	59	100			
CDPK-E	43	41	41	36	100		
CDPK-F	57	59	39	98	58	100	
CDPK-G	65	70	60	43	53	53	100

The *M. polymorpha* CDPK sequences are fairly diverse in terms of homology. Protein identity ranges from 36.6 to 97.9 % (Table.3.2) and protein similarity ranges from 57.7 to 98.9 % (Table.3.3). The similarity score demonstrates that many of the substitutions of amino acids are for chemically similar amino acids. Ranges for similar structure usually cut off at 40 % identity (Rost 1999). However, the CDPKs of *Arabidopsis thaliana* have a similar protein identity range of 36.5-94.9

Table 3.3: The protein sequence similarity of the 7 putative CDPKs from *M. polymorpha* from pairwise alignment using EBLOSUM62 matrix.

	CDPK-A	**CDPK-B**	**CDPK-C**	**CDPK-D**	**CDPK-E**	**CDPK-F**	**CDPK-G**
CDPK-A	100						
CDPK-B	87	100					
CDPK-C	82	69	100				
CDPK-D	74	75	71	100			
CDPK-E	63	58	59	53	100		
CDPK-F	74	74	71	99	58	100	
CDPK-G	80	84	76	70	62	70	100

% (Table.3.3). Therefore the protein sequences of *M.polymorpha* have a similar diversity in sequence to those from *A. thaliana*.

The phylogeny may be tentatively used to propose functions of each *M. polymorpha* CDPK based on literature of the most similar CDPKs. Subgroup I contains MpCDPK-A/B/G which may imply a function in quite a few areas as this groups seems particularly divergent (Fig.3.1). AtCPK5 and AtCPK6 are part of subgroup I involved in upregulation of WRKY transcription factors which respond to a range of biotic and abiotic stresses (Boudsocq et al. 2010; Boudsocq et al. 2012; Munemasa et al. 2010). In *A. thaliana* subgroup I are also AtCPK1/2/4/11 which are known to phosphorylate RBOHD and NADPH oxidase which are involved in oxidative burst as part of the hypersensitive response. AtCPK6 is also part of subgroup I and involved in methyl-jasmonate and abscisic acid signalling in stomatal guard cells as part of plant immunity, drought, and salt stress responses (Xing et al. 2001; Boudsocq et al. 2010; Munemasa et al. 2010; Hubbard et al. 2012). Therefore, it may be likely that one of these proteins may be involved in salt stress or in plant immunity.

MpCDPK-C was the CDPK identified by Nishiyama et al. (1999) and therefore both splice variants have been included in the phylogeny in subgroup II (Fig. 3.1) but were so similar that they have been included as a single branch. Interestingly, MpCDPK-C is most closely related to AtCPK17 and AtCPK34 which are involved in pollen tube formation (Myers et al. 2009). However MpCDPK-C is also similar to AtCPK3 has a range of proposed functions, including during herbivory, salinity and

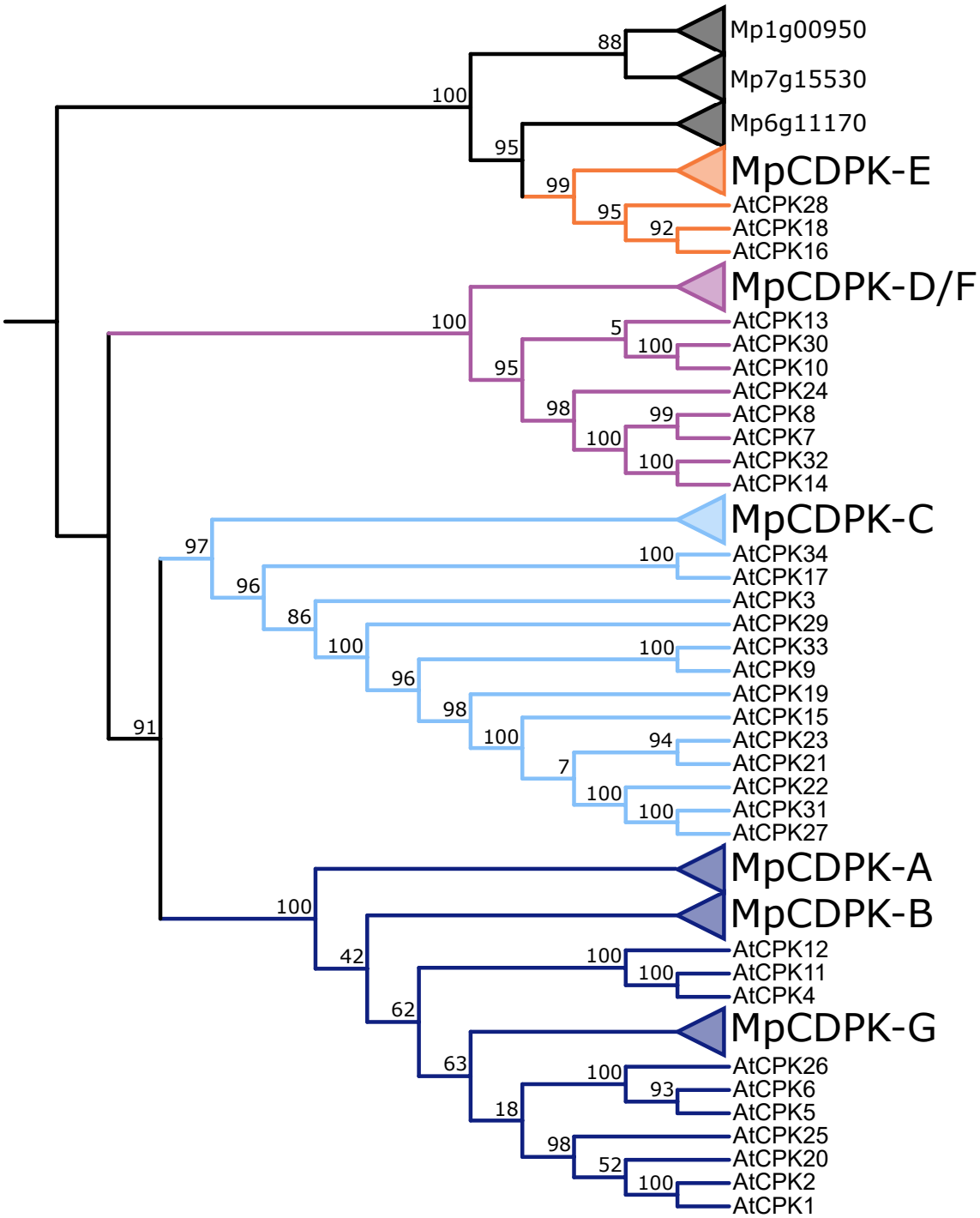


Figure 3.1: A phylogenetic tree of the CDPKs from *M. polymorpha*. Phylogeny was generated by a nearest neighbour alignment (Mafft, 100 iterations, 100 retrees) with sequences from *A. thaliana* for comparison

stomatal closure, and is also reported as soluble, plasma membrane targeted, and tonoplast localised which may imply it is involved in a range of context dependent processes (Boudsocq et al. 2012; Hubbard et al. 2012; Boudsocq and Sheen 2013). This pair of spliceoforms seems to be differentially expressed in the sexual structures

of the plant and the only difference between them is an alternative exon 6 used. This exon encodes a region including only EF hand 2 which could imply differential calcium binding between the two variants. Therefore, it may be that a mutant in MpCDPK-C would have some form of developmental phenotype for the sexual structures or potentially some phenotype involved in the fertilisation process and actually be more similar to AtCDPK17/34 than AtCPK3.

MpCDPK-D was the CDPK described by Sharma et al. (2013) (Fig.3.1). MpCDPK-D and its male counterpart MpCDPK-F are in subgroup III and groups with AtCPK30 and AtCPK10. AtCPK30/10/32 activate NIN-like transcription factors involved in nitrogen uptake and seem responsive to ABA and therefore sensitivity to both factors will be interesting to investigate in *M. polymorpha* knockouts of MpCDPK-D/F (Liu et al. 2017). Both MpCDPK-D and MpCDPK-F are the only MpCDPKs found in our analysis to harbour no introns. Current hypotheses for the emergence of intronless genes centre around them being generated through whole genome duplication events or retrotransposition events from cDNAs (Liu et al. 2020; Tutar 2012). As *M. polymorpha* is not thought to have gone through any whole genome duplications retrotransposition is more likely; which often results in the gene becoming a pseudogene due to a lack of 5' promoter and 3' polyadenylation site. However the description of this gene by Sharma et al. (2013) demonstrates that it is both expressed and changes expression throughout the lifecycle implying its transcription is controlled and it is therefore not a pseudogene. Liu et al. (2020) found a correlation between intronless genes and activity in salt and drought responsive pathways in *Arabidopsis* and rice which, combined with homology to AtCPK10 and AtCPK30, may imply a ABA responsive pathway function.

MpCDPK-E is the only other CDPK and it is found in subgroup IV (Fig.3.1). This subgroup contains AtCDPK28 which has been characterised to functionally interact with BIK1 in immune signalling and is the only CDPK known to bind CaM (Bender et al. 2017). MtCDPK1 also falls into subgroup IV and is involved with downregulating immune signalling in symbiosis, this may imply that MpCDPK-

E is involved with adaptive immunity or plant-microbe interactions. However the rest of the proteins in this subgroup have no proposed functions yet, and therefore MpCDPK-E may have different properties.

Table 3.4: Characteristics of CDPKs from *M. polymorpha* includings predicted localisation, intron number and subgroup. PM is an abbreviation of Plasma Membrane.

CDPK	**Localisation**	**Intron Number**	**Subgroup**
CDPK-A	Nucleus	5	I
CDPK-B	Nucleus	6	I
CDPK-C	Nucleus	6	II
CDPK-C2	Nucleus	6	II
CDPK-D	Nucleus	0	III
CDPK-E	Nucleus/PM	10	IV
CDPK-F	Nucleus	0	III
CDPK-G	Nucleus	5	I

The intron structure of the CDPKs can be indicative of subgroup and function. *A. thaliana* subgroups I/II have similar average introns number (6.7, 6.25) and range (5-9, 6-8) of introns. CDPK subgroup III has a higher average number of introns (7.1) but a similar range to subgroup II (6-8). Subgroup IV has distinctly more introns with an average of 11.7 and range of 11-13. The intron structure of *M. polymorpha* CDPKs reflects the same trends in *A. thaliana*. In subgroup I MpCDPK-A/B/G have 5/6/5 introns respectively. Subgroup II MpCDPK-C isoforms both have 6 introns (Table.3.4). Therefore in these subgroups the average number of introns seems lower than *A. thaliana* but similar. Subgroup III MpCDPK-D/F both have no introns and are located on the sex determining chromosomes which is different from *A. thaliana*. While sex-linkage is not seen in *A. thaliana* CDPKs it has be observed in *Papaver rhoeas* (Rudd et al. 1996) and the CDPK functions in determining self-incompatibility and pollen tube growth, and in *Silene latifolia* (Nichloas et al. 2005). MpCDPK-E is the only subgroup IV representative and has 10 introns which is similar but fewer than the 11-13 introns in *A. thaliana*.

Table 3.5: Characteristics of CBLs from *M. polymorpha* includings IDs, length and intron structure.

CBL	V3.1	V5.1	Length (aa)	Length (bp)	Intron
CBL-A	Mapoly0015s0061	Mp2g07750	232	696	7
CBL-B	Mapoly0066s0053	Mp4g00900	247	741	8
CBL-C	Mapoly0134s0040	Mp5g19810	213	639	7

3.2.2 There are 3 CBLs in *M. polymorpha*

Using the FamilyFinder script, 3 gene IDs of potential CBLs were identified in *M. polymorpha* (Table.3.5). The sequences retrieved were then scanned with Expasy Prosite (De Castro et al. 2006). All of the sequences had detectable EF hands for EF2, EF3 and EF4. EF1 could not be detected in *A. thaliana* or *M. polymorpha* sequences by Prosite due to its noncanonical 14 amino acid loop region. However the sequence can be manually found due to the set distances between EF hands in CBLs and confirmed based on similarity to other CBLs which all have the 14 amino acid loop for EF1.

Some differences were found between the *M. polymorpha* and *A. thaliana* CBLs in the EF hands. The canonical EF hand in *A. thaliana* is twelve amino acids with a consensus DxDxDGxxDxxE sequence (Kolukisaoglu et al. (2004); Haswell and Meyerowitz (2006); Li et al. (2009)). Amino acids in positions 1(X), 3(Y), 5(Z), 7(-X), 9(-Y) and 12(-Z) are particularly important in conservation for calcium coordination but variation can take place in other positions resulting in changes in calcium binding affinity (Sánchez-Barrena et al. 2013). A two amino acid insertion between position X and Y resulted in the formation of the fourteen amino acid CBL EF1. The fourteen amino acid EF hand in MpCBL-A matched the consensus sequence of *A. thaliana* (SxxxIxDGLxNxxE) but EF1 of the other MpCBLs show some differences (Larebeke et al. 1974). Both MpCBL-B and MpCBL-C show a mutation of the conserved N to H which is a change from a polar uncharged amino acid to a positively charged amino acid in position -Y. MpCBL-C also has a change in position Y from V to I which are both hydrophobic side chains. Even though there are these differences from *A. thaliana* CBLs, the proposed CBLs found in *M.*

polymorpha still have the domains required for function and therefore can still be defined as true CBLs. However the calcium binding potential of the first EF hand in the MpCBLs are unknown. Edel and Kudla (2015) also detected 3 CBLs using version 3.1 of the genome and the FamilyFinder script could find no more MpCBLs in version 5.1 of the genome (marchantia.info).

Table 3.6: The protein sequence identity of the 3 putative CBLs from *M. polymorpha* from pairwise alignment using EMBOSS Water.

	CBL-A	**CBL-B**	**CBL-C**
CBL-A	100		
CBL-B	66	100	
CBL-C	68	85	100

Table 3.7: The protein sequence similarity of the 3 putative CBLs from *M. polymorpha* from pairwise alignment using EBLOSUM62 matrix.

	CBL-A	**CBL-B**	**CBL-C**
CBL-A	100		
CBL-B	89	100	
CBL-C	87	93	100

All of the CBLs in *M. polymorpha* share a high protein identity demonstrating that they have been correctly identified by FamilyFinder. The shared protein identity is above 60 % for the three genes. MpCBL-B/C show the highest protein identity of 85.4 percentage identity shared (Table.3.6). Proteins similarity between the CBLs is even higher at 86.9-93.4 % (Table.3.7). High protein similarity is expected as CBLs generally have a very conserved structure with specific distances between EF hands and conservation between EF hands. The CBL proteins also have a conserved FPSF phosphorylation site present in all the CBLs of *M. polymorpha*. The main difference between CBLs tends to be in the N-terminal variable domain that confers localisation and is either the MGCxxS/T motif for palmitoylation and myristoylation which is conserved in MpCBL-B/C or a tonoplast localisation motif. As this is the most variable domain in CBLs, the phylogeny tends to group CBLs by localisation.

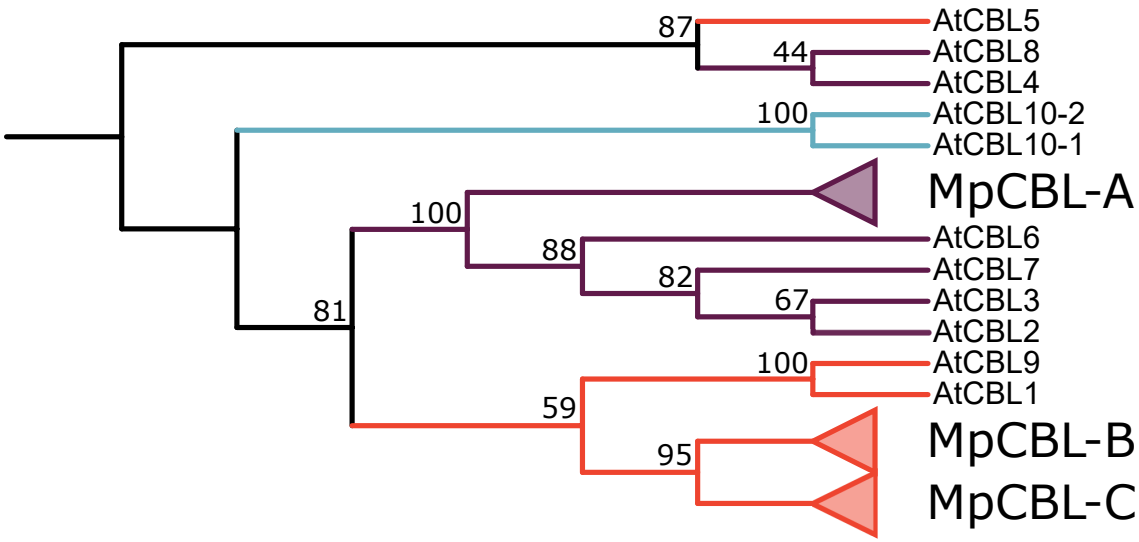


Figure 3.2: A phylogenetic tree of the CBLs from *M. polymorpha*. Phylogeny was generated by a nearest neighbour alignment (Mafft, 100 iterations, 100 retrees) with sequences from *A. thaliana* for comparison

Subcellular localisation is often proposed for CBLs based on phylogeny but function can also be suggested as localisation is linked to function. The MGCxxS/T motif can be identified from the N-terminus of the protein. The tonoplast localisation sequence has been defined in based on alignment of protein sequence multiple CBLs which localise to the tonoplast. However, the homology does not always indicate localisation as can be seen in a CBL in *P. taeda* (PITA000027276) which is lipid modified and plasma membrane localised despite falling into the subgroup of AtCBL10 which contains a transmembrane sequence usually associated with plasma membrane targetting (Edel and Kudla 2015). Therefore, the potential localisation of *M. polymorpha* CBLs were proposed here and matched the previous work of Edel and Kudla (2015). CBL-A showed phylogenetic clustering with the tonoplast targeting group of CBLs including the known tonoplast localising *A. thaliana* CBLs (AtCBL-2/3/6/7) but did not contain the three cysteine residues in its first twenty amino acids which were proposed to confer tonoplast localisation by Batistič et al. (2010) (Fig. 3.2). The MpCBL-A sequence also did not contain the tonoplast targeting sequence consensus (MSQCxDGxKHxCxSxxxCF) proposed by Tang et al. (2012) and Kleist et al. (2014). Therefore, it could be that localisation machinery is different between *M. polymorpha* and *A. thaliana* for tonoplast localisation, and

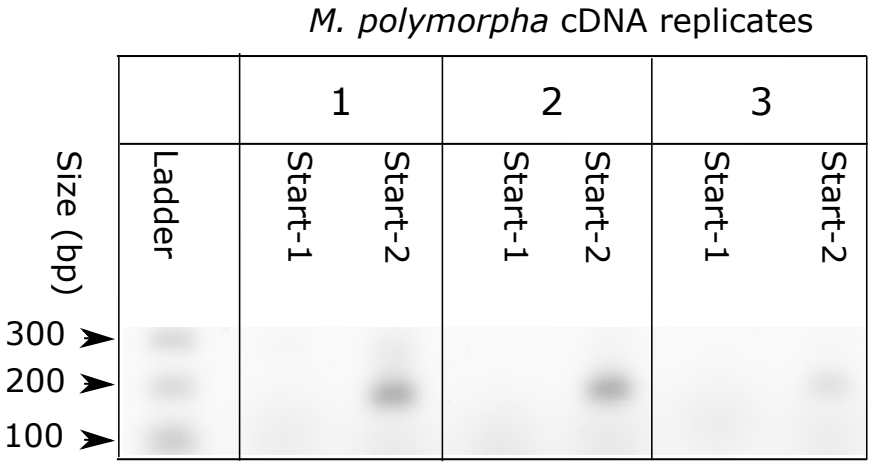


Figure 3.3: Identification of the CBL-B start-site. The CBL-B start site was assessed by PCR with primers for the start based on the predicted start from the genome (Start-1) or the start of the MGCxxS/T motif (Start-2) for 3 replicates of cDNA extracted from *M. polymorpha*. The electrophoresis was carried out on a 1% gel with a 1 Kb plus ladder for size comparison.

the same targeting motifs are not used. Alternatively it may not localise to the tonoplast and has a much more similar protein sequence in terms of other domains including EF hand sequence.

In contrast to MpCBL-A, MpCBL-B and MpCBL-C cluster in the phylogenetic tree with AtCBL1 and AtCBL9 which are plasma membrane targeted (Fig.3.2; Batistič et al. (2010); Weinl and Kudla (2009)). Upon investigation, both proteins contain the MGCxxS/T motif proposed to confer plasma membrane targeting by lipid modification through myristoylation and S-acylation (Batistič et al. 2008). However, the targeting motif does not seem to be located on the N-terminus of the protein in MpCBL-B unlike AtCBL-1/9 or MpCBL-C which may be due to the methods in which the protein sequence was determined. The difference could also be from difference in localisation in *M. polymorpha* as the predicted start-site is instead found 34 amino acids N-terminal of the MGCxxS/T motif (Mao et al. 2016). Expression of this gene would need to be tested to find if it is expressed from this N-terminal methionine or if it is actually expressed from the MGCxxS/T motif.

To correctly identify the start site of MpCBL-B, cDNA was extracted from *M. polymorpha* plants for 3 replicates and PCR was used with primers designed for the computationally predicted start site on the genome (V.5.1, marchantia.info), or for

the MGCxxS/T motif start-site. The genome listed start-site was designated start-1 and the MGCxxS/T motif potential start-site was listed as start-2. Only start-2 could be amplified with the primers across 3 biological replicates from the cDNA which indicates that the start-site is from the MGCxxS/T motif (Fig.3.3). This would be expected if the motif is functional as myristoylation and palmitoylation takes place the the N-terminus of the protein. The implication from this is that CBL-B is localised to the plasma membrane through the MGCxxS/T motif and therefore all phylogenies presented in this thesis use the MGCxxS/T motif as the start-site for the protein.

While there is not a known link between CBL intron number and function or subgroup, some differences have been noted in intron number in different species. The MpCBLs contain the same range of introns (7-8) as those from *A. thaliana* (7-8) and *Oryza sativa* (7-8). Typically only CBLs with a transmembrane domain have 8 introns and all other CBLs have 7 introns in *A. thaliana* and *O. sativa* (Kolukisaoglu et al. 2004) and yet MpCBL-B has 8 introns with the MGCxxS/T motif instead (Table.3.5). *MpCBL-A* and *MpCBL-C* have the same intron number (7) despite *MpCBL-A* being more different in term of phylogeny than *MpCBL-B* and *MpCBL-C* (Table.3.5). *Physcomitrella patens* has a slightly different intron structure to both *M. polymorpha* and *A.thaliana* with 6-7 introns per CBL with two intronless CBLs (Liu et al. 2020).

3.2.3 There are 2 CIPKs in *M. polymorpha*

Table 3.8: Characteristics of CIPKs from *M. polymorpha* includings IDs, length and intron structure.

CBL	**V3.1**	**V5.1**	**Length (aa)**	**Length (bp)**	**Intron**
CIPK-A	Mapoly0005s0039	Mp1g05680	441	1326	13
CIPK-B	Mapoly0025s0017	Mp2g26670	443	1332	13

Using the FamilyFinder script; 9 gene IDs of potential CIPKS were identified in *M. polymorpha*. The sequences were then scanned with Prosite (De Castro et al. 2006) for a detectable kinase domain and NAF domain required for interaction with

CBLs (Table.3.8). Only two had both of the required domains and were labelled CIPK-A and CIPK-B. The MpCIPKs have a high shared identity (71.2%) and similarity (86.1%). Due to the reduced number of CIPKs in *M. polymorpha* it is difficult to propose function. Furthermore many CIPKs interact with more than one CBL to carry out multiple different functions such as AtCIPK23 which is involved in potassium and nitrate uptake with AtCBL1/9, or iron uptake through interactions with AtCBL2/3 (Rodenas and Vert 2021). The presence of only 3 detectable CBLs and 2 CIPKs means that it is likely that more than one interaction will also be possible for the MpCIPKs perhaps also for multiple functions. Therefore to deduce potential function the interactions will need to be tested between the MpCIPKs and MpCBLs.

However the CIPKs present still share relation to some CIPKs of known function in *A. thaliana*. MpCIPK-A falls basal to the group containing MpCIPK-B and AtCIPK8/24 therefore it may be basal to this group or an early diverging group similar to AtCIPK26/3/23/9. If CIPK-A is in the early diverging group it may be function like AtCIPK23 which is known for positive regulation of AKT1 for potassium uptake via phosphorylation. Also of this group is AtCIPK26 is known to be involved in regulation of NAPH oxidases in *A. thaliana* and ABA-regulated ubiquitin ligase Keep on Going (KEG; Lyzenga et al. (2017); Drerup et al. (2013)). AtCIPK3 is also linked to ABA signalling for drought, cold, saline and wounding responses (Pandey et al. (2008), Sanyal et al. (2017b)). Similarly, AtCIPK9 is associted with potassium uptake in low potassium soil as well as various ionic and drought stress responses (Liu et al. 2013). CIPK-B is most closely related to AtCIPK8/24. AtCIPK24 is more widely known as Salt Overly Sensitive 2 (SOS2) for its known regulation of SOS1 which is a Na^+/H^+ antiporter and involved in salt tolerance (Ji et al. 2013). AtCIPK8 has recently been described as an alternative SOS CIPK with regulation by AtCBL10 for control of SOS1 (Yin et al. 2019). All of the AtCIPKs in this group with ascribed function are involved in ionic responses such as Na^+ and K^+ balance and drought responses and therefore it may be proposed that the MpCIPKs are

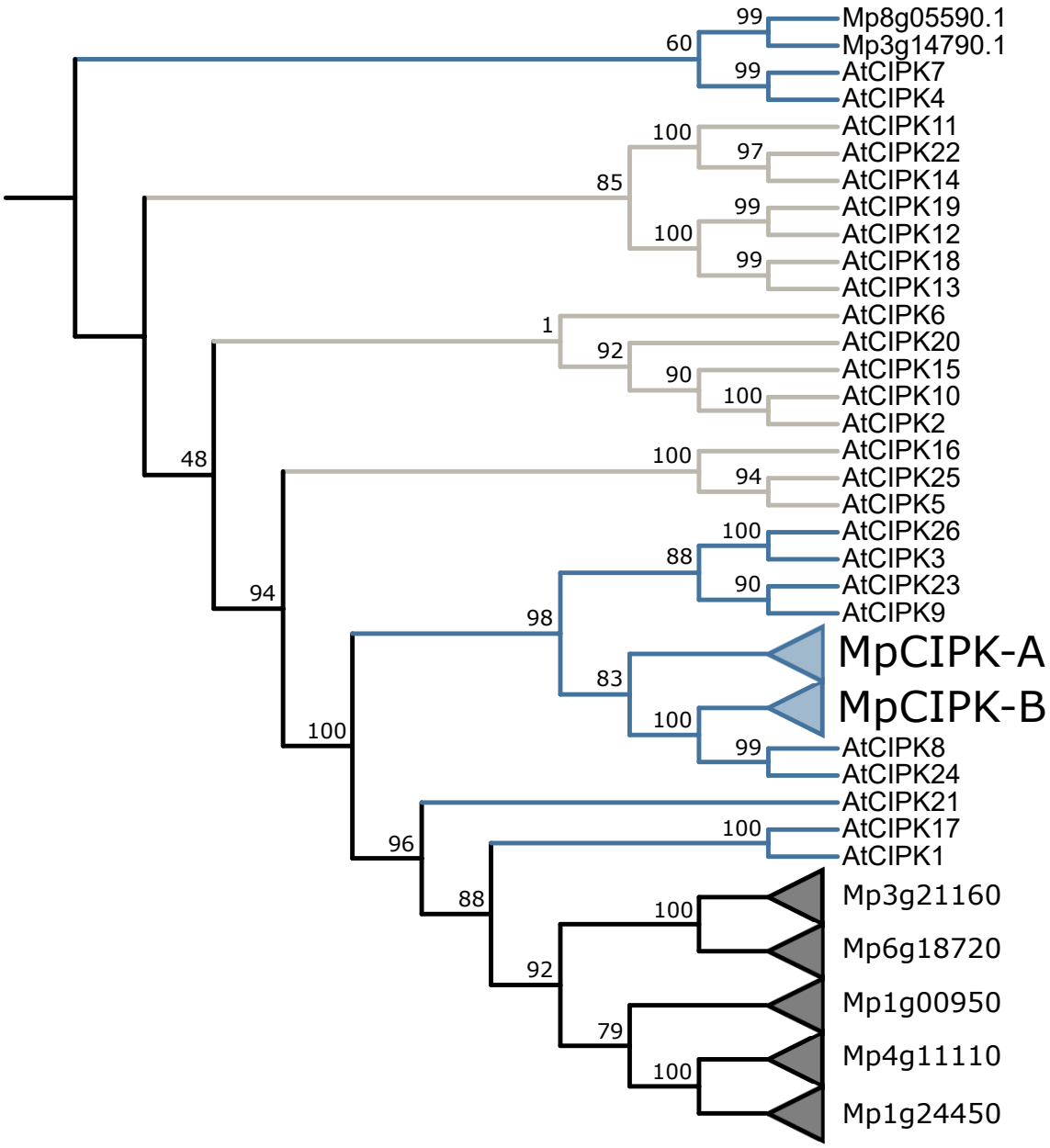


Figure 3.4: A phylogenetic tree of the CIPKs from *M. polymorpha*. The phylogeny was generated by a nearest neighbour alignment (Mafft, 100 iterations, 100 retrees) with sequences from *A. thaliana* for comparison

likely involved in these responses.

CIPKs are defined split into four clades: clades I-III are intron-rich (>8) and clade IV is intron-poor (<3). The only detectable CIPKs using FamilyFinder on v5.1 and v3.1 of the genomes of *M. polymorpha* were both part of the intron-rich clade of CIPKs which confirms the result of Edel and Kudla (2015). The lack of intron-poor CIPKs may be indicative of loss of this clade or that *M. polymorpha* diverged from the green lineage prior to the evolution of this clade. Previously it

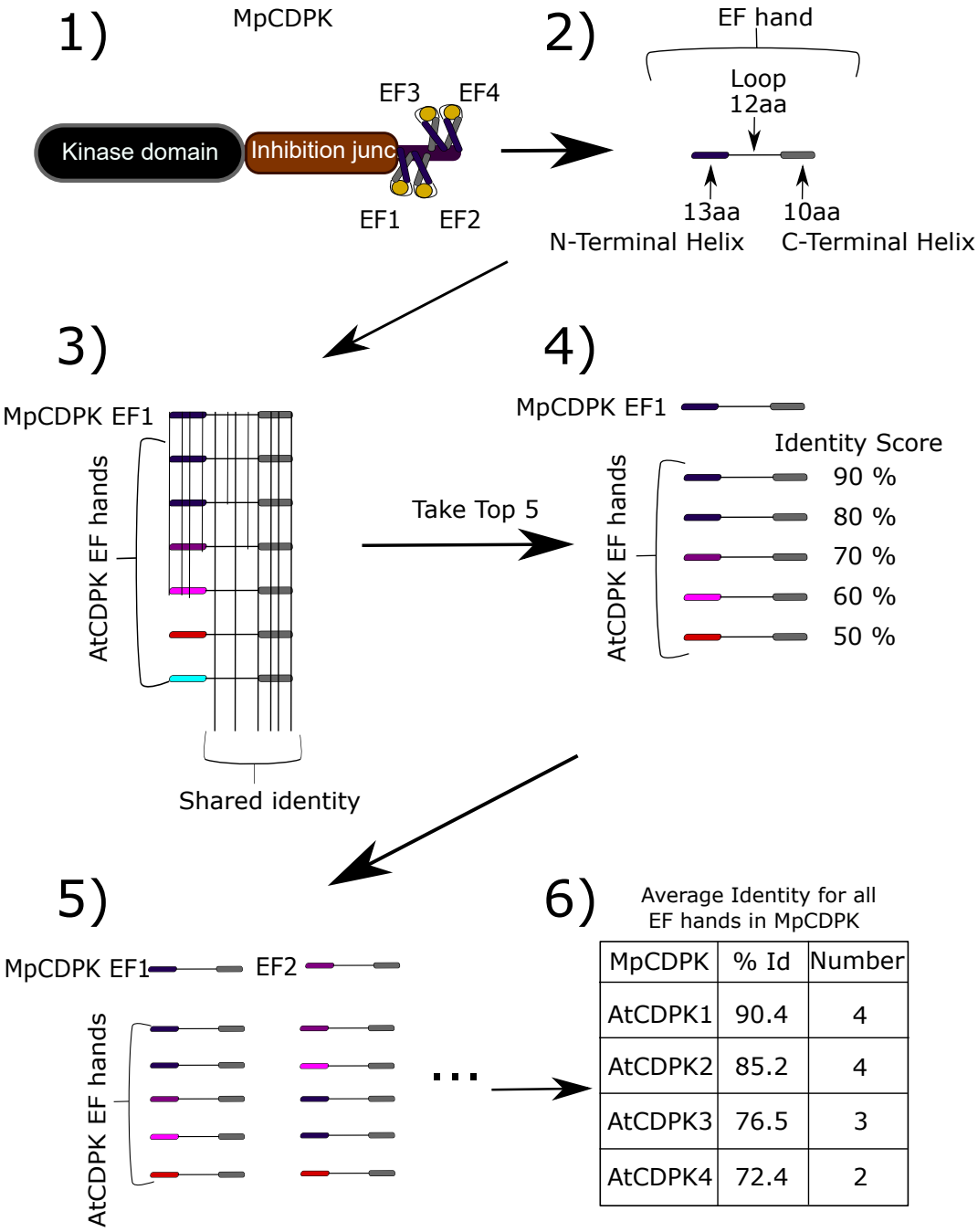
has been proposed that the loss of introns took place in the seed-plants resulting in neofunctionalisation of the CIPKs for improved drought tolerance (Zhu et al. 2016).

3.2.4 EF hand Analysis

Due to the conserved nature of EF hands, comparison of the EF hands may indicate function better than phylogeny of the full protein sequence. The analysis took the EF hands from *M. polymorpha* and *A. thaliana* CDPKs and CBLs including helices as identified by Prosite scan (Fig.3.5.2). If the helices could not be identified, 13 aa N-terminal and 10 aa C-terminal of the loop were taken (Fig.3.5.2). The first EF hand of each CBL used the 14 amino acid loops and added 13 aa N-terminal and 10 aa C-terminal of the extended loop. Protein identity matrices were generated by Clustal Omega (Fig.3.5.3). Each EF hand from an *M. polymorpha* calcium decoder had the five highest conserved identity assessed (Fig.3.5.4). The first EF hand was then compared to the other EF hands (EF2+EF3+EF4) in the same protein in *M. polymorpha* to see if the EF hands in *A. thaliana* were conserved at each EF hand position (Fig.3.5.5). An average identity score was generated across the four EF hands of each protein to find the highest scoring *A. thaliana* calcium decoder to each *M. polymorpha* calcium decoder (Fig.3.5.6). The analysis should confirm the subgroup of each calcium decoder assessed and provide a single calcium decoder from *A. thaliana* that has the highest EF hand identity to the calcium decoder from *M. polymorpha*.

CDPK EF hands confirm each subgroup

MpCDPK-A and MpCDPK-B are both in subgroup I of the CDPKs and closely related to AtCPK2. AtCPK2 was also found to be most similar to both of these proteins from the EF hand analysis. Each EF hand in the MpCDPK-A had AtCPK2 in its top 5 most similar EF hands and it had a average identity across these 4 EF hands of 70.7 %. The first 3 EF hands of MpCDPK-B has a high average identity of 79.6 % with the first 3 EF hands of AtCPK2. The fourth EF hand of AtCPK2



MpCDPK	% Id	Number
AtCDPK1	90.4	4
AtCDPK2	85.2	4
AtCDPK3	76.5	3
AtCDPK4	72.4	2

Figure 3.5: Schematic of the EF hand analysis. 1) EF hands for each calcium decoding protein from N- to C-terminus from 1-4. 2) EF hand sequences were retrieved and defined as the loop sequence with 13 amino acids N-terminal and 10 amino acids C-terminal to include both helices of the EF hand (identified by Prosite or manually). 3) identity was calculated for each EF hand of each calcium decoder from *M. polymorpha* compared to each other EF hand from *A.thaliana* (using Clustal Omega). 4) the top 5 hits were taken for each EF hand of each calcium decoder from *M. polymorpha* along with the identity score of each of those hits. 5) each of the EF hands from each calcium decoder is included and the average identity was calculated. 6) the highest average identity was calculated for each calcium decoder by including the top 5 hits for each EF hand from that calcium decoder.

was not found in the top 5 highest identity for MpCDPK-B. AtCPK1 also had 3 EF hands in the top 5 of MpCDPK-B with an average identity of 75.81 % with the last 3 EF hands of MpCDPK-B. AtCPK1 and AtCPK2 both fall into subgroup I of *A. thaliana* which was the same subgroup that the phylogeny implied for MpCDPK-A/B (Fig.3.1) and therefore the EF hand analysis confirmed the subgroup. However neither AtCPK1/2 were considered the most similar to MpCDPK-A/B by phylogeny but both AtCPK1/2 are involved in NADPH oxidase activity, and therefore it will be interesting to see if MpCDPK-A/B are involved in these processes.

MpCDPK-C is the only *M. polymorpha* CDPK in subgroup III and is closely related to AtCDPK34 and AtCDPK17 by phylogeny (Fig.3.1). Similarly the EF hand comparison also showed a high average similarity in the EF hands to both AtCPK17 (74.4 %) and AtCPK34 (73.6 %). The EF hand analysis therefore confirmed the subgroup III allocation of MpCDPK-C and also confirmed the most closely related CDPKs from *A. thaliana* were AtCPK17/34. As MpCDPK-C isoforms are known to be differentially expressed in sexual structures which provides some implication that the analysis may accurately predict involvement in sexual structures. However more investigation is required to know if it is involved in fertilisation similar to AtCPK17/34.

MpCDPK-D and MpCDPK-F are both in subgroup II and share 97.9 % identity with each other and 98.9 % similarity. EF hands 1 and 4 have 100 % identity between MpCDPK-D and MpCDPK-F and 97.2 % identity between EF hands 2 and 3 so both proteins likely respond to very similar calcium signals. However the proteins are encoded on the sex chromosome and therefore they are seperate in adult plants as single sex chromosomes confer sex in haploid *M. polymorpha*. However due to these difference they have higher similarity to different proteins. MpCDPK-D has a top 5 hit with AtCPK13 for 3 EF hands with average identity of 64.8 % . MpCDPK-F only has 3 EF hands in the top 5 for AtCPK30 with and average identity of 64.8 %. However despite the diference in most identical EF hands. MpCDPK-D still has a very high identity with AtCPK30 and has 3 EF hands in the top 5 with an

average identity of 63.89 %. The same is not true for CDPK-F as it does not have AtCPK13 in the top 5 EF hands with conserved identity. AtCPK13/30 are both in subgroup III with MpCDPK-D/F (Fig.3.1) which confirmed the ascribed subgroup. AtCPK30 activate NIN-Like transcription factors and respond to ABA (Liu et al. 2017) and AtCPK13 inhibits KAT1/2 shaker K^+ channels and reduce stomatal opening (Ronzier et al. 2014). Therefore it will be interesting to look for responses involved in drought as both AtCPK13/30 are involved in drought responses.

MpCDPK-E is the only *M. polymorpha* CDPK in subgroup IV and is most similar by phylogeny to AtCPK28 (Fig.3.1). The EF hand analysis found 4 EF hands in common with AtCPK28, AtCPK16 and AtCPK18 of subgroup 4 with the highest average protein identity with AtCPK16 of 71.1 %. The EF hand analysis therefore confirms subgroup IV grouping of MpCDPK-E. AtCPK16 is involved in maintainence of salt induced calcium signals by phosphorylating AtGLR3.7 which generates the nucleus localised salt induced calcium signal (Wang et al. 2019). This indicates that MpCDPK-E may also be involved in salt induced calcium signalling. The involvment in salt induced calcium signalling is interesting as it implies that, despite the diverse subgroups present in MpCDPKs, many of them may have involvment in salt or drought signalling through diverse mechanisms.

CBL EF hand cannot be used to confirm subgroup

Using a protein identity matrix for comparison on EF hands is not effective for comparison of the CBL EF hands when only the *A. thaliana* EF hands are included as all MpCBLs shared the highest conserved EF hand identity with AtCBL2/3. MpCBL-A had the highest protein identity with 4 EF hands in common with AtCBL-2/3 and 80.7%/81.4% identity respectively. MpCBL-B has highest identity with AtCBL-2 with 4 EF hands in common and 74.6% identity. MpCBL-C has 4 EF hands in common with AtCBL-2/3 with 71.1% shared identity with both. However MpCBL-B/C are not predicted to localise to the tonoplast like AtCBL2/3 and likely do not truly respond to the same signals as AtCBL2/3. It could be that all the CBLs from

M. polymorpha have similar EF hands and therefore there is no difference between subgroups which would make EF hand analysis ineffective for CBLs. However, it is more likely that protein sequence identity is insufficient to analyse the differences between EF hands and perhaps similarity is also required to compare EF hands effectively for CBLs. It could also be that only including the 10 *A. thaliana* EF hands does not add the required resolution for this kind of analysis and therefore both will be required for a more in depth analysis.

Table 3.9: The First EF-hand of the MpCBLs compared to SOS3 from *Arabidopsis*. CBL EF1 loop sequence is marked with the canonical 12 amino acid loop sequence with calcium coordinating residues marked with their geometry as X, Y, Z, -Y, -X, -Z. The inserted 2 amino acids in the 14 aa loop are marked with a '-'.

CBL	**X**	**-**	**-**	**Y**	**3**	**Z**	**5**	**6**	**-Y**	**8**	**-X**	**10**	**11**	**-Z**
CBL-A	S	S	A	V	I	D	D	G	L	I	N	K	E	E
CBL-B	S	S	T	V	I	D	D	G	L	I	H	K	E	E
CBL-C	S	S	T	V	V	D	D	G	L	I	H	K	E	E
SOS3	S	S	S	I	I	D	D	G	L	I	H	K	E	E

Both MpCIPKs seem to fall basal to AtCIPK8/24 which have both been demonstrated to be involved in salt tolerance responses to calcium signals through phosphorylation regulation of the SOS1 ion channel. Therefore paying particular attention to the first EF hand of the CBLs may allow us to discover a SOS3 homologue in MpCBLs. The C-terminus of the loop of the EF hand domain is known to tolerate less mutations and both MpCBL-B/C have an identical N-terminus loop to SOS3 including a conserved DGLIHKEE with 8 amino acids identical of the total 14 amino acids. The N-terminus has more variation and both MpCBL-B/C third residue have a S>T substitution in the third position and a I>V substitution in the fourth position. The fifth amino acid of the 14 amino acid loop is different between the MpCBLs and B has a conserved I with SOS3 and MpCBL-C has a I>V difference which are both branched chain amino acids. Based on this single fifth position amino acid difference between MpCBL-B/C, MpCBL-B is the most similar to SOS3.

3.3 Discussion

3.3.1 Proposed MpCDPK function based on bioinformatic analysis

There are 7 CDPKs in *M. polymorpha* with representatives of each subgroup known in later-diverging plants such as angiosperms (Fig.3.6). There are four subgroups of CDPKs known in higher plants which is typically linked to function, however the function can be vague as not all CDPKs have ascribed function due to functional redundancy. CDPK-A/B/G were all identified as part of subgroup I which is typically associated with salt and drought responses. CDPK-C was the only CDPK in *M. polymorpha* in subgroup II which is associated with pollen tube formation through AtCPK17/AtCPK34. In fact, CDPK-C labelled in this study has already been described to have alternative spilceoforms which are differentially expressed in the sexual structures of *M. polymorpha* (Nishiyama et al. 1999). CDPK-D/F are encoded on the V and U sex chromosomes respectively and therefore are not expressed together in haploid adult plants of *M. polymorpha* and form the only representative of subgroup III. The AtCPKs that are part of subgroup III typically are involved in drought or salt responses such as AtCPK10/30/32 but also a range of biotic stress responses including AtCPK10/30/14/32. Subgroup IV representative CDPK-E is in a subgroup with AtCPK16/18/28 which have functions in development (AtCPK16), or biotic stress responses to viruses (AtCPK18/28) and bacterial stress (AtCPK28) and is likely involved in developmental or biotic stress responses (references).

In higher plants subgroups II and III are further divided into groups 'a' and 'b' but this is not present in *M. polymorpha*. *M. polymorpha* subgroups II and III have single representatives in an adult plant and therefore cannot be subdivided into 'a' and 'b' and it seems that both of the representatives in these groups fall basal to the whole group in comparison to *A. thaliana*. This finding is unsurprising as similar comparative studies have found that the same is true of other bryophytes *P. patens* and lycophytes *S. moellendorfii* which implies the expansion and subdivision of these

subgroups took place later in the green lineage (Valmonte et al. 2014). Therefore both CDPK-C and CDPKs-D/F may be interesting to investigate as potentially they retain the original functions of these subgroups. This may be of paticular interest for CDPK-D/F as they form very similar homologues (97.9% identity) found on different sex chromosomes of *M. polymorpha* and also have no introns which may imply that they newly arose as a subgroup in *M. polymorpha*. Also CDPK-C could be of interest but with its spliceoforms linked to differential expression in sexual structures it may have developmental phenotypes which could be difficult to characterise.

EF hand analysis of the CDPKs in *M. polymorpha* may imply which candidate is closest in relation to a CPK in *A. thaliana*. CDPK-A/B have a high conservation of EF hands with AtCPK-1/2 and also 87.2 % identity with each other which may imply functional redundancy with each other. So while they likely operate in similar functions in response to drought and salt stress they may not be the best candidates to investigate further due to the risk of functional redundancy. CDPK-C had highest conserved EF hands with AtCPK17 (74.4 % identity) and then with AtCPK34 (73.6 %) averaged across all four EF hands. CDPK-C likely operate in a similar manner as both AtCPK17/34 are involved in fertilisation and pollen tube formation. While CDPK-D/F have 97.9 % identity with each other it seems that they have some differences in EF hand 2 and EF hand 3 with 97.2 % identity. It may be interesting to investigate the differences in EF hands in CDPK-D/F cause any differences in calcium binding dynamics and the signals to which they respond. Of the *A. thaliana* CPKs AtCIPK13 had the highest EF indentity conserved with CDPK-D (63.9 %) and AtCPK30 had the highest conserved EF hand identity with CDPK-F (64.8 %). The difference in the EF hand analysis are expected when the main area of difference between the proteins is in the EF hands but it may imply they have different calcium signal response dynamics. CDPK-E had highest EF hand identity shared with AtCPK16 (71.1 %) which functions in development. This type of analysis is not typical and it will be interesting to see if the proposed closest candidates based on EF hand match in terms of function as it could add some resolution to proposing

function based purely on overall phylogeny by instead proposing function based on calcium binding domains.

3.3.2 Proposed MpCBL and MpCIPK function based on bioinformatic analysis

There are 3 confirmed CBLs in *M. polymorpha* as were described by Edel and Kudla (2015) using version 5.1 of the genome (Fig.3.6). The CBLs confirmed in this study were confirmed by version 5.1 of the genome and therefore the most up to date version of the genome currently available. Typically localisation can be determined by investigating the first 20 amino acids of the protein sequence. Either proteins have plasma membrane localisation due to a MGCxxS/T motif that confers posttranslational modification sites for myristoylation and palmitoylation or a tonoplast localisation sequence conferred by either 4 cystienes in the first 20 amino acids (Batistič et al. 2010) or MSQCxDGxKHxCxSxxxCF motif (Mao et al. 2016). Phylogeny implied tonoplast localisation in CBL-A but neither of the tonoplast localisation sequences could be seen in the first 20 amino acids and therefore a different mechanism may be present that confers the localisation or its localisation is not to the tonoplast. CBL-B/C were both proposed to localise to the plasma membrane by phylogeny (Fig. 3.2) and the start-site was confirmed as the MGCxxS/T motif for MpCBL-B/C (Fig. 3.3). Therefore MpCBL-B/C localise to the plasma membrane and MpCBL-A is propose to localise to the tonoplast but this is not confirmed.

There are two confirmed CIPKs in *M. polymorpha* as were described by Edel and Kudla (2015) using version 5.1 of the genome (Fig.3.6). Function and intron number can be proposed and checked using phylogeny for CIPKs. All of the CIPKs in *M .polymorpha* fall into the intron-rich clade as would be expected as intronless CIPKs seemingly arose in angiosperms (Zhu et al. 2016). They also all contain multiple introns as would be expected of this clade. Based on phylogeny CIPK-B seems to be most closely related to AtCIPK8 or AtCIPK24 (SOS2). AtCIPK8 is thought to be involved in drought stress responses and AtCIPK-24 was one of the

first defined CIPKs involved in salt tolerance through phosphorylation control of SOS1 an Na^+/H^+ antiporter after activation by SOS3 (AtCBL4). CIPK-A falls as an outgroup to CIPK-B, AtCIPK8 and AtCIPK24 and therefore it is likely that the early function of the CBL-CIPK network is in drought and salt responses.

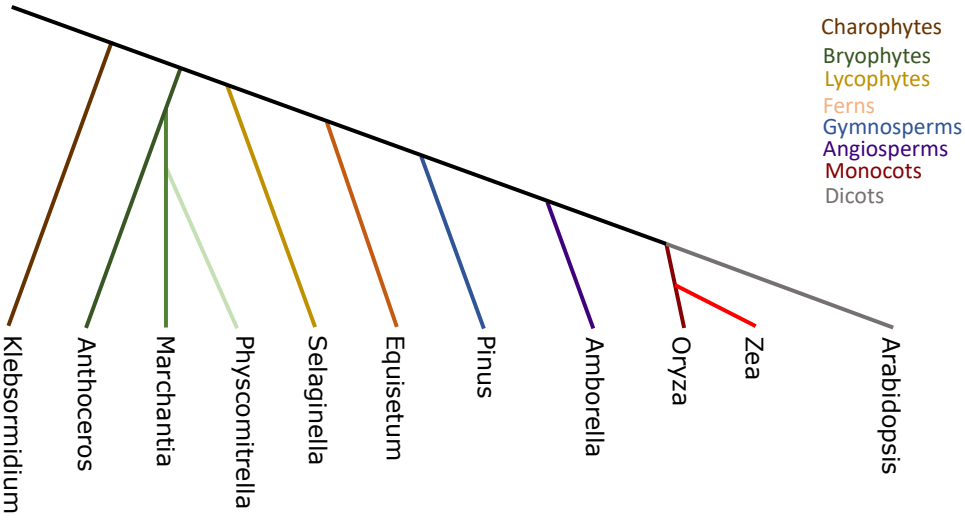
As the CIPKs seem to be predominantly involved in drought and salt stress responses the EF hand analysis was also carried out on the CBLs on *M. polymorpha*. The analysis of the CBL EF hands from *M. polymorpha* to those from *A. thaliana* only demonstrated similarity on average with AtCBL2/3 which fall into the tonoplast localising subgroups. While it may be that CBL-A is similar to the AtCBL2/3 and has a similar function it is unlikely that CBL-B and CBL-C share function with their predicted localisation to the plasma membrane. Hence the EF hand analysis may not function for CBLs or it may be that only including the 10 AtCBLs doesnt provide adequate resolution. Direct comparison of the first EF hand of SOS3 (AtCBL4) which is the known interactor of AtCIPK24, and the better characterised of AtCIPK8/24, demonstrated a high similarity between CBL-B/C and SOS3 which highest similarity between CBL-B and SOS3.

As salt and drought conditions are common functions for CBLs/CIPKs and CDPKs this may be the best phenotype to investigate and determine a function of a calcium decoding gene in *M. polymorpha*. So far only a single *P. patens* CIPK has a proposed function and PpCIPK1 has been functions in salt tolerance (Xiao et al. 2021). While PpCIPK1 was not included in the phylogeny presented in this thesis, it seems to be basal to both AtCIPK24 and AtCIPK23 clades but not part of the AtCIPK21 clade (Fig.3.4) (Kleist et al. 2014). Alongside this, SOS1 has been identified in *P. patens* with two different homologues. It may be that the two homologues of SOS1 in *P. patens* is due to the the recent genome dulplication it has been through (Fraile-Escanciano et al. 2010). The whole genome duplication is given also given as a potential reason for why CBLs seem to be high in number in *P. patens* and form cognate pairs (Kleist et al. 2014). Therefore it may not be true that *M. polymorpha* has two copies of SOS1 due to the lack of genome duplication although

the genome implies a wide diversity of ion channels present (Bowman et al. 2017). The analysis of CIPKs and CBLs here demonstrates the link between both CIPKs to potential drought and salt responsive functions a therefore it would be an obvious phenotype to test by knockouts of these genes. The MpCDPKs also seem to show high homology to proteins involved in salt and drought responses. However, the AtCPKs they show homology to are predominantly involved in guard cell control of stomata and it is currently not known that *M. polymorpha* air-pores can control conduction (Ruszala et al. 2011; Ishizaki et al. 2013b; Jones and Dolan 2017). However the link between calcium decoding proteins and salt or drought tolerance would be interesting to investigate as a reason for why this group of signalling proteins expanded with the emergence onto land. With the emergence onto land, dessication, high salinity and cold are some of the most important stresses plants would have to deal with. It would be fascinating if it was found the calcium decoders original functions could be to help provide a level of tolerance to these stresses.

3.3.3 There is little functional redundancy in *M. polymorpha* calcium decoders based on bioinformatic analysis

More broadly *M. polymorpha* has already shown features that imply it has more basal mechanisms compared to higher plants in a number of signalling pathways. For example MpDELLA has been identified and can still respond to oxidative stress and other stress conditions but completely independent of the complexities of gibberellic acid signalling (Hernández-García et al. 2021). *M. polymorpha* can also still respond to ethylene but cannot produce ethylene itself and instead uses the ethylene precursor, 1-aminocyclopropane-1-carboxylic acid (ACC), and utilises ACC in a similar manner to ethylene signalling found in angiosperms (Li et al. 2020). *M. polymorpha* also cannot synthesise jasmonate-isoleucine and instead uses jasmonate precursor, OPDA (Monte et al. 2019). While it may be true that not all of these signalling examples represent ancient signalling pathways and actually may be loss of function similar to how stomata were lost in *M. polymorpha* (Harris et al. 2020). However



CBL	3	?	3	4	4	?	7	4	11	12	10
CIPK	1	?	2	7	5	?	26	12	34	43	26
CDPK	12	?	7	21	10	?	13	17	29	41	34

Figure 3.6: Phylogeny of plant and algae species with known calcium decoders show a general increase in the calcium decoding toolkit. Genus is shown for described species: *Klebsormidium flaccidum, Anthoceros augustus, Marchantia polymorpha, Physcomitrella patens, Selaginella moellendorfii, Pinus taeda, Amborella trichopoda, Oryza sativa, Zea mays, Arabidopsis thaliana. M. polymorpha* decoder numbers proposed from this chapter, all other calcium decoder numbers from Edel and Kudla (2015). Equisetum and Anthoceros remain cryptic groups for calcium decoders. Phylogeny of plants and green algae proposed by Puttick et al. (2018).

this can only be elucidated by comparative phylogenomics across the bryophytes and may still provide indications of how the green lineage may have survived and thrived without all of the complex signalling processes found in angiosperms. The same concept can also be applied to calcium signalling and calcium decoding and comparative phylogenomics and functional analysis of genes and proteins involved in this signalling can provide insight into the core requirements of calcium signalling and its evolution.

Based on the bioinformatic analysis of the sequences there is little indication of extensive functional redundancy in calcium decoding proteins of CDPK, CBL and CIPK families (Fig.3.6). CDPKs may have the highest evidence for potential functional redundancy but only in subgroup I as 3 CDPKs are present in *M. polymorpha* in this subgroup. Therefore *M. polymorpha* offers previously unprescedented potential to understand the function of its calcium decoding toolkit as it may be possible to

carry out genetic knockout of all of the calcium decoding genes and ascribe function. *M. polymorpha* as a model can therefore be used to deeply interrogate the mechanistic action of these different calcium decoding protein families. *M. polymorpha* also provides opportunity to develop evolutionary insight into how these different protein families have expanded gone through neofunctionalisation across the green lineage through macroevolutionary comparisons with much more well studied plants such as *A. thaliana*. Similarly comparisons to other bryophytes with much wider expansion of the calcium decoding network such as *P. patens* may provide insights to key steps in the evolutionary process, such as genome duplication, and how this process has affected calcium decoding proteins in particular.

Chapter 4

Expression Analysis of Calcium Decoding Genes of *M. polymorpha* in Response to Salt Stress

4.1 Salt Tolerance is an Economically Important Abiotic Stress Tolerance which is Mediated by Calcium Signals

Salinization of arable land is a leading threat to sustainable global food security and plants demonstrate reduced yield at concentrations as low as 40 mM NaCl (Munns 2005). It is estimated that 33% of irrigated land globally is afflicted by high salinity and the salinized areas are increasing by 10% a year and it is predicted to reach more than 50% of arabale land by 2050 (Jamil et al. 2011). This threat of salinization is likely to be amplified by climate change through rising sea levels (Chen and Mueller 2018).This will result in a reduction in global food production as it stands and the negative effects of climate change would only be exacerbated further if total land usage increases. In the last 50 years agricultural land usage has increased by 9% and yield has more than doubled (Pretty 2007). Therefore it has been possible to

increase yield without massively increasing land usage and this methodology is the most sustainable approach to increasing global food security. To do this, constant improvement in understanding of plant physiological responses to salt and tolerance mechanisms that already exist in plants are required.

4.1.1 Salt Stress is a Multifactorial Stress Containing both Ionic and Osmotic components

Osmotic stress is one component of salt stress responses due to the changes in water potential in the soil caused by high salinity. Many of the salt tolerance responses include these osmotic stress responses due to the osmotic component of salt stress. For example, of the 167 genes expressed in response to 250 mM NaCl in *Arabidopsis thaliana*, 101 were also induced by drought treatment (Seki et al. 2002). However, the types of responses vary across time and can be divided into early and late responses. Early osmotic stress responses deal with immediate issues caused by drought stress such as shutting off transpirational flow by closing stomata. Included in this, is the induction genes responsible for production of compatible solutes or osmoprotectant proteins such as Late Embryogenesis Abundant 1 (LEA1). Similarly in both osmotic and salt stress MKK4 is involved in the accumulation of reactive oxygen species (ROS) and MPK3 which induces biosynthesis of absicsic acid (ABA) through NCED3 (Kim et al. 2011). ABA and H_2O_2 both function in drought tolerance, one mechanism for this is through stomatal closure and limiting water loss due to transpiration (Sirichandra et al. 2009). All of these responses occur very quickly upon a high salt stimulus. Longer term effects of drought, that can also be found under salt stimuli, include development of a smaller leaf area which are often thicker due to changes in cell size and shape (James et al. 2002). Other changes in response to drought and salt stress include differential rates of root elongation and lateral root initiation (Rubinigg et al. 2004). Therefore many of the drought responses found in plants also take place in response to salt.

The core difference between osmotic stress and salt stress is the ionic component

which the plant has to resolve with salt specific conditions. Generally the accumulation of Na^+ is a particular issue due to the negative electrochemical potential gradient and chemical concentration gradient which encourages Na^+ uptake into the plant (Demidchik et al. 2002). Previous studies on *Triticum turgidum* subspecies demonstrated that plants with reduced sodium concentrations after 10 days of 150 mM NaCl treatment have a higher grain yield under saline conditions (Husain et al. 2003). This phenotype has been associated with effective salt extrusion at the root to exclude 98% of the salt from the transpiration stream (Lindsay et al. 2004). Similarly the Na^+ can be compartmentalised to the vacuole to reduce potentially damaging effects of ion toxicity in the cytoplasm through action of vacuolar Na^+/H^+ antiporters such as NHX1 (Blumwald and Apse 2000). However, K^+ and organic solutes are required to accumulate in the cytoplasm and organelles to balance the osmotic effects of Na^+ (Flowers and Yeo 1986). Therefore the ionic effects of salt focus on compartmentalising or extruding Na^+ and recomposing osmotic potential through off-setting with organic solutes or nutritionally important ionic equipoise through K^+/Na^+ balance.

4.1.2 Calcium Signals respond to Salt Stress Systemically in Plants

Calcium signals take place in different subcellular and tissue locations in response to salt. Some of the original detections of calcium signals in response to salt were characterised using aequorin reaching peaks of 1.5 μM cytosolic calcium within 25 seconds in *A. thaliana* (Knight et al. 1997). The limit of these studies so far is that they cannot differentiate whether the origin of cytosolic calcium release is from extracellular spaces or from the vacuole. However these signals travel systemically and it has been defined that salt stress induced calcium waves can travel systemically at 400 μm/s and rely on vacuolar ion channel TPC1 (Choi et al. 2014). The reliance on TPC1, which is localised to the vacuole, could imply that the cytosolic calcium release stimulated by salt is from the vacuole. Alternatively it could be that the

systemic signal relies on initial vacuolar release, but perhaps there is also a seperate release from extracellular spaces and further resolution could still be added to this area.

The salt induced calcium signal can vary between species and within species. Previous studies have demonstrated that in *A. thaliana* different accessions can have distict signals that vary based on accession. For example, Col-0 produces two peaks with the first peak taking place in the first 25 seconds and the second between 30 and 120 seconds but C24 accession does not have a second peak (Schmöckel et al. 2015). Since then it has been confirmed that the two distinct intracellular signals in Col-0 in response to salt have different subcellular localisations; one in the nucleus and the other in the cytosol (Huang et al. 2017). In both of these accessions of *A. thaliana* the amount of NaCl applied increases the amplitude of the peak to higher concentrations of Ca^{2+}. *N. benthamiana* has a biphasic peak similar to Col-0 but under increasing NaCl concentration there are different changes in amplitute in the first and second step of the peak which is referred to as the peak ratio (Mithöfer and Mazars 2002). However the only other plant that has been transformed with a calcium reporter and treated with salt is *Oryza sativa* but low resolution was used in monitoring the difference and therefore it is not possible to detect any differences in calcium peaks in this plant (Zhang et al. 2015). Further studies are required to fully elucidate the variety of signals that can take place in response to single stimuli between species and between accessions.

Previous studies in our lab have identified a salt responsive calcium signal in *M. polymorpha*. Tak1 accession of *M. polymorpha* was transformed with a R-GECO fluorescent reporter for calcium signals under a constitutive promoter and were treated with 100 mM NaCl or 200 mM sorbitol (Fig.4.1). Under NaCl treatment a biphasic peak was detected which continued over 40 minutes post treatment. The salt induced signal was not present when pre-treated with lanthanum chloride which is a nonselective calcium ion channel blocker mainly effective on cytosolic calcium ion channels, however as TPC1 responds to cytosolic calcium it likely also blocks cal-

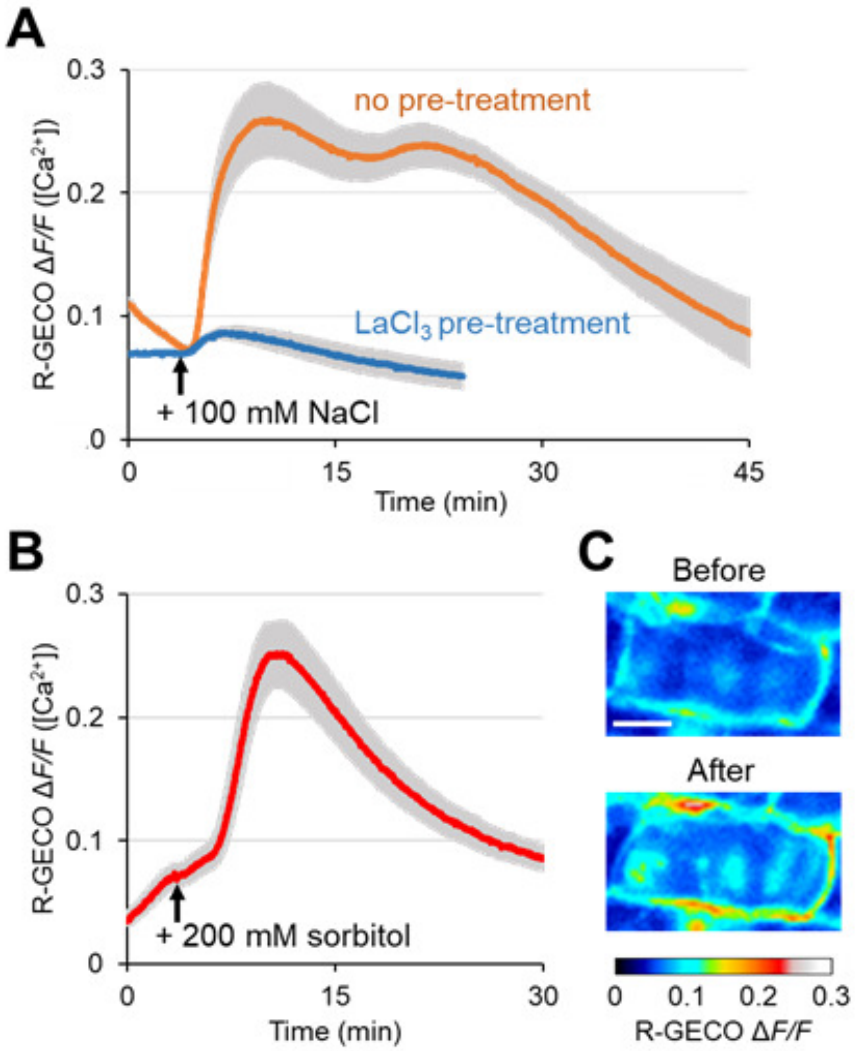


Figure 4.1: Calcium Signals detected in R-GECO Tak1 in response to Salt and Sorbitol. A) R-GECO Tak1 *M. polymorpha* detected a biphasic signal in response to salt (100 mM NaCl) over 40 minutes that could not be detected with lanthanum chloride pretreatment. B) R-GECO Tak1 *M. polymorpha* detected a monophasic signal in response to osmotic stress (200 mM sorbitol) over 30 minutes. C) Fluorescence images of the salt induced signal before and after salt treatment. White bar represents 10 μm. (Image provided by Dr Ben Miller, Unpublished)

cium flux through this channel as well (De Vriese et al. (2018), Fig.4.1). Conversely sorbitol had a single peak which took place over 30 minutes (Fig.4.1). Generally it seems that both signals took much longer than the detected signals in *A. thaliana*, *N. benthamiana* or *O. sativa* taking place over many minutes rather than the two minutes for most salt signals in other plants. However the signal from sorbitol and NaCl seem to be distinct due to the monophasic versus biphasic signal. The distinct subcellular localisations of the signal have not been investigated, nor have systemic

signals in *M. polymorpha*.

4.1.3 Calcium Decoders in Response to Salt Stress

Calcium Dependent Protein Kinases can directly improve salt tolerance by reducing ROS production or by maintaining calcium signalling over time in response to salt. *O. sativa CPK12* over expression lines have a higher tolerance to salt stress and mutants in this gene are more sensitive. Part of the reason for this tolerance was found to be related to lower ROS production, this may be because the over expression line of *OsCPK12* had reduced expression of *OsrbohI* which is an NADPH oxidase and increased expression of *OsAPx2* and *OsAPx8* which are ROS scavenging enzymes (Asano et al. 2012). ROS production has a direct effect on salinity tolerance, however the overexpression lines were found to also have increased susceptibility to blast fungus as the plants could not generate sufficient ROS to defend against the pathogen, which makes it undesirable as a biotechnological target to increase salt tolerance. *A. thaliana* CPK16/3/34 have also been found as potential targets that directly affect salt tolerance through maintainence of salt induced calcium signals. GLR3.7 is a calcium ion channel that generates salt induced calcium signals and interacts with At14-3-3w protein. When the salt induced signal is generated these CPKs promote salt tolerance through phosphorylating the 14-3-3w binding site on AtGLR3.7 (Ser860) thereby abolishing interaction with 14-3-3w (Wang et al. 2019). Both of these methods show direct activity on ROS production and control of the salt induced calcium signal to increase salt tolerance through the activity of CDPKs.

CDPKs can also influence salt tolerance through defence against osmotic stress more generally often through control of ABA influenced pathways and reducing stomatal conductance. AtCPK6 is thought to be involved in salt tolerance through osmotic control by generating compatible solutes to reduce the effects of water potential stress. Overexpression of *AtCPK6* results in a more tolerant plant and higher proline and malondialdehyde. However, knockouts are not more susceptible to salt implying there may be some functional redundancy (Xu et al. 2010b). AtCPK10 is

also involved in salt defence as a drought tolerance CPK through interaction with heat shock protein 1 to control stomatal conduction as part of ABA and Ca^{2+} signalling (Zou et al. 2010). Similarly, *AtCPK4/11* mutants show pleiotropic ABA insensitivity phenotypes including seed germination, seedling growth and stomatal movement. AtCPK4/11 has involvement in osmotic effects involved in ABA signalling responses through phosphorylation of ABA-responsive transcription factors ABF1/4 which leads to these diverse ABA associated phenotypes (Zhu et al. 2007).

The CBLs that have been characterised so far in *A. thaliana* that respond to salt induced calcium signals directly confer salt tolerance through extrusion of Na^+ or through rebalancing ions through K^+ uptake. The Salt Overly Sensitive pathway (SOS) was the first CBL-CIPK pathway described in plants and description of that pathway can be found in the introduction (1.3.1). AtCBL2/3 are vacuole localised CBLs that interact with V-ATPases for pH balancing which also affects ionic sequestration. Similarly AtCBL1/9/4 interact with AKT1/2 shaker-like channels to control K^+ uptake for nutrient balancing and can act to counter the K^+ depletion caused by increasing Na^+ concentration under high salinity (Tang et al. (2012);Li et al. (2006); Lee et al. (2007); Held et al. (2011); Nieves-Cordones et al. (2012)). However CIPKs can also be involved in both salt and osmotic effects. SOS2 (AtCIPK24) is stimulated by salt induced calcium signals through SOS3 (AtCBL4) and AtCBL10 but it similarly interacts with protein phosphotase 2C ABI2. In this system SOS2 is dephosphorylated by ABI2 and knockout of ABI2 increases salt tolerance to salt shock and abscisic acid insensitivity (Ohta et al. 2003).

4.1.4 Calcium Decoder Expression Changes in Response to Salt Stress

AtCPKs can act as positive and negative regulators of salt tolerance networks in *A. thaliana*. Typically both positive and negative regulators are upregulated in response to both osmotic and salt stress in the first 32 hours but negative regulators are downregulated after the 32 hour time-point, and only positive regulators remain increased

in expression under continuous salt stress. AtCPK12 is a positive regulator of salt tolerance and is upregulated in response to salt and controls ROS production to help coordinate salt tolerance. Generation of RNAi lines demonstrated that knockdown interferes with CPK12 mediated salt tolerance and plants are more susceptible to salt (Zhang et al. 2018). *AtCPK4/11* are involved in the osmotic part of salt stress tolerance through ABA pathways and increases in expression allow protein products to mediate the osmotic component of salt tolerance responses. However, negative regulator CPKs of osmotic stress tolerance can also be upregulated in response to salt, such as *AtCPK9* or *AtCPK33* in *A. thaliana* (Gong et al. 2005; Li et al. 2016). However it is likely because *AtCPK9/33* are upregulated in response to ABA, as a negative regulator of stomatal closure, and ABA increases under drought and salt tolerance (Chen et al. 2018). *AtCPK33* is actually downregulated by the 24 hour timepoint. The studies of expression of *AtCPK9* took place at relatively short time points from 0-32 hours it is also likely downregulated at timpoints after 32 hours as would be expected under long term salt stress. The reason for upregulation of genes associated with reduced salt tolerance (hence negaative regulators) is likely so that the salt tolerance network can quickly downregulate salt tolerance mechanisms if the salt that triggers the response is not continuous.

CBL expression is more complicated in response to salt as it is more tissue specific, likely as it acts directly on ionic effects, and therefore gene expression increases in tissues are affected by ionic stress in that specific tissue. Specifically in the Salt Overly Sensitive pathway, *AtCBL4 (SOS3)* is upregulated in response to salt in the cortex and endodermis but not the epidermis in response to salt after 1 hour. However, the ion channel *SOS1* is upregulated in the epidermis and therefore it may be that high expression of *SOS1* is to export Na^+ to the soil and that instead of increasing expression of *SOS1* in the endodermis and cortex *AtCBL4 (SOS3)* is upregulated for rapid control of SOS1 activity in those locations. Thereby increasing SOS1 activity in response to Na^+ inclusion, if Na^+ concentration surpasses the capacity of SOS1 to export Na^+ in the epidermis (Ji et al. 2013). Studies in

Vitis vinifera demonstrate that *SOS3.7/.8* homologues are specifically upregulated by salt in the root but are downregulated in response to osmotic stress (Ma et al. 2019). A model could be proposed where AtSOS3 is activated by calcium signals under osmotic stress and salt stress, and as Na^+ concentration must be maintained under osmotic stress, *AtSOS3* is downregulated to stop erronous export of Na^+. However, when Na^+ concentration is also high *AtSOS3* is upregulated to allow it to increase activity of SOS1. Expression of *AtCBL10* is specifically in the stem and shoot and it operates in salt tolerance through interaction with SOS2 to mediate SOS1 activity. *AtCBL10* seems to have a slower increase in expression peaking at 24 h, likely due to the activity of AtSOS3 blocking salt inclusion into the stem and shoot to a later time point (Quan et al. 2007). The *SOS* genes are all downregulated at 6 days in reponse to salt stress which indicates that SOS may be an early salt responsive signalling system (Rolly et al. 2020). Therefore the SOS system can be considered an early responsive system to salt in *A. thaliana* that initially increases in expression in a tissue specific manner in the first 24 h, and decreases in expression sometime after 24 h, and is fully downregulated by 6 days.

4.1.5 Are there Salt Responsive Calcium Decoders in *Marchantia polymorpha*?

To understand the salt responsive calcium decoders it is important to determine what level of tolerance *M. polymorpha* has to salt stress. Two different accessions are available for *M. polymorpha*; the Takaragaike accession from Japan and the Cambridge accession which may have different tolerance due to the different environments from which they were obtained. Previously *M. inflexa* has been found to have different dehydration tolerance in the different sexes of the plant (Marks et al. 2016) and therefore it is important we also assess if this is true and applies to salt tolerance in *M. polymorpha*. If differences are found it may be interesting to see if there are any differences in calcium decoders between the accessions and sexes.

To be able to determine the salt responsive function of some of the calcium

decoders expression analysis can be used in *M. polymorpha*. As a number of the CBLs, CIPKs and CDPKs respond to osmotic and salt stress, and due to the economic importance of these stresses salt stress, expression can be analysed in a dosage dependent manner to determine which of these decoders may be involved in this stress response. As CDPK negative and positive regulators of response to salt both increase in expression in early responses to salt and negative regulators being downregulated at later time points. The best timepoint to check for expression of CDPKs seems to be post day 6, so that positive regulators are upregulated and negative regulators downregulated in response to salt stress. Similarly, CBLs may be upregulated initially in *A. thaliana* response to salt, but are downregulated at later timepoints, and therefore expression can also be assessed post day 6 for downregulation in *M. polymorpha*. As salt responsive calcium decoder can change in expression in multiple plant tissues (1.3.1), and as do not currently know which tissues calcium decoders may be expressed in *M. polymorpha*, using the whole plant for RNA extraction seems more appropriate to capture the expression changes of all the CDPKs, CBLs and CIPKs.

4.2 Results

4.2.1 Tak1 *M. polymorpha* Survives from 0-150 mM Salt for 10 days

To assess the tolerance dosage of salt that *M. polymorpha* could survive and determine its tolerance to salt 10 cuttings of 1 month old Tak1 *M.polymorpha* were placed on 1/2 MS media plates (1% sucrose) plates containing 0, 50, 100, 150, 200 mM NaCl and observed after 10 days for survival. Tak1 plants demonstrated stress responses from 50 mM salt with visible anthocynanin production in response to stress. From 100 mM chlorosis is detectable with bleaching in some of the cuttings but survival will still possible at this concentration for 10 days. At 200 mM NaCl all of the plants are bleached from earlier than 10 days. Active growth seemed reduced

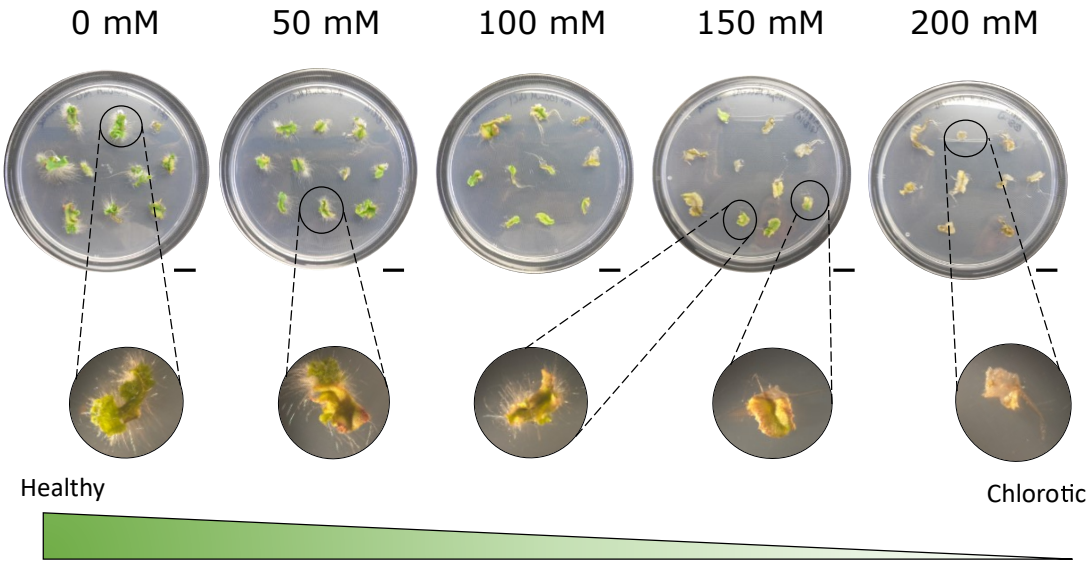


Figure 4.2: Tak1 *M. polymorpha* can survive 150 mM salt for 10 days.10 Cuttings of *M. polymorpha* Tak1 are treated with 0-200 mM NaCl and are increasingly chlorotic at higher concentrations of salt when treated for 10 days. Closer images of some selected cuttings are beneath to demonstrate chlorotic response. Black bars represent 1 cm.

in salt conditions compared to 0 mM salt concentrations (4.2).

4.2.2 Tak1 *M. polymorpha* is More Sensitive to Salt than Sorbitol

To have consistent *M. polymorpha* plantlets to test, gemmae were taken and grown under non stress conditions on 1/2 MS and sucrose (1%) for 7 days and then transferred to stress plates containing 0, 50, 100, 150, 200 mM for 7 days. As plantlets were smaller on 50-150 mM NaCl stress plates growth seems to be suppressed. As a control for Cl^- stress, and as a comparison between KCl and NaCl, plantlets were also placed on KCl stress plates from 50-150 mM and demonstrated a similar reduced growth phenotype. *M. polymorpha* grown under the same conditions and transferred to sorbitol to test for the effect of osmotic stress. Double the concentration of sorbitol compared to NaCl/KCl is known to have the same reduction in water potential in the plate (e.g. 100 mM sorbitol or 50 mM NaCl) and hence causes an equal osmotic stress effect without the ionic effects of stress. 100-400 mM sorbitol have the same reduction in growth but do not show the same chlorosis as plants under NaCl

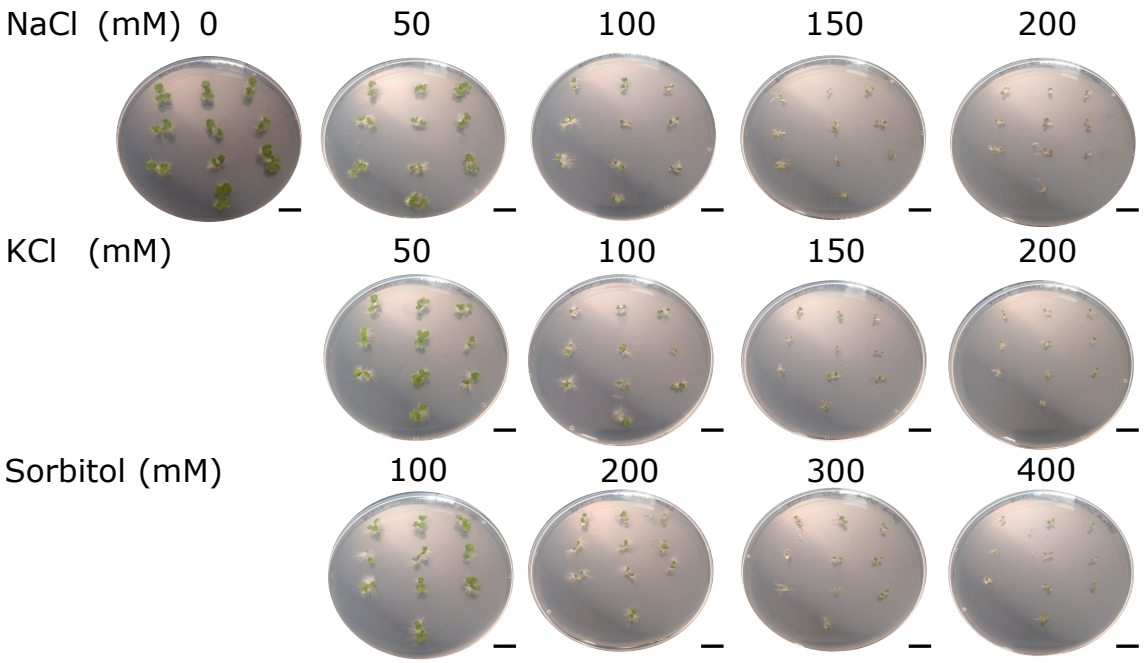


Figure 4.3: Tak1 *M. polymorpha* gemmae can survive 150 mM NaCl or KCl and 200 mM Sorbitol. *M. polymorpha* Tak1 Gemmae grown for 1 week and then transferred to various plate treatments from 0-200 mM NaCl, 0-200 mM KCl and isoosmotic concentrations of sorbitol from 0-400 mM. Growth deficiency is visible after 7 days of treatment for 10 plantlets at higher concentrations but plantlets survive better on sorbitol than on ionic stresses of KCl and NaCl. Black bars represent 1 cm.

or KCl stress (4.3). Reduced chlorosis under sorbitol conditions is expected as while they have the same osmotic pressure but do not have the ionic affects of NaCl or KCl stress. Interestingly, gemmae grown for 1 week do not seem to produce visible anthocyanin but do show stunted growth under NaCl stress. Also under NaCl stress some green tissue was still observable at 200 mM NaCl which meant that some of the gemmae dormancy affect may still be present in such young plants (Eklund et al. 2018). If gemmae dormancy, which is controlled by ABA response, does effect the salt stress response in *M. polymorpha* it may also effect expression of calcium decoding genes. When gemmae are grown to a later timepoint, hence avoiding the dormancy, there is much more variety in growth prior to salt treatment (data not shown). A lack of consistency in growth of starting material would affect the ability to measure growth reductions caused by salt. Therefore, to test adult salt tolerance of *M. polymorpha* it was decided to use consistent cuttings from adult plants for comparison.

4.2.3 Tak and Cam Accessions of *M. polymorpha* Demonstrate Different Growth Rates on 1/2 MS and Sucrose (1%)

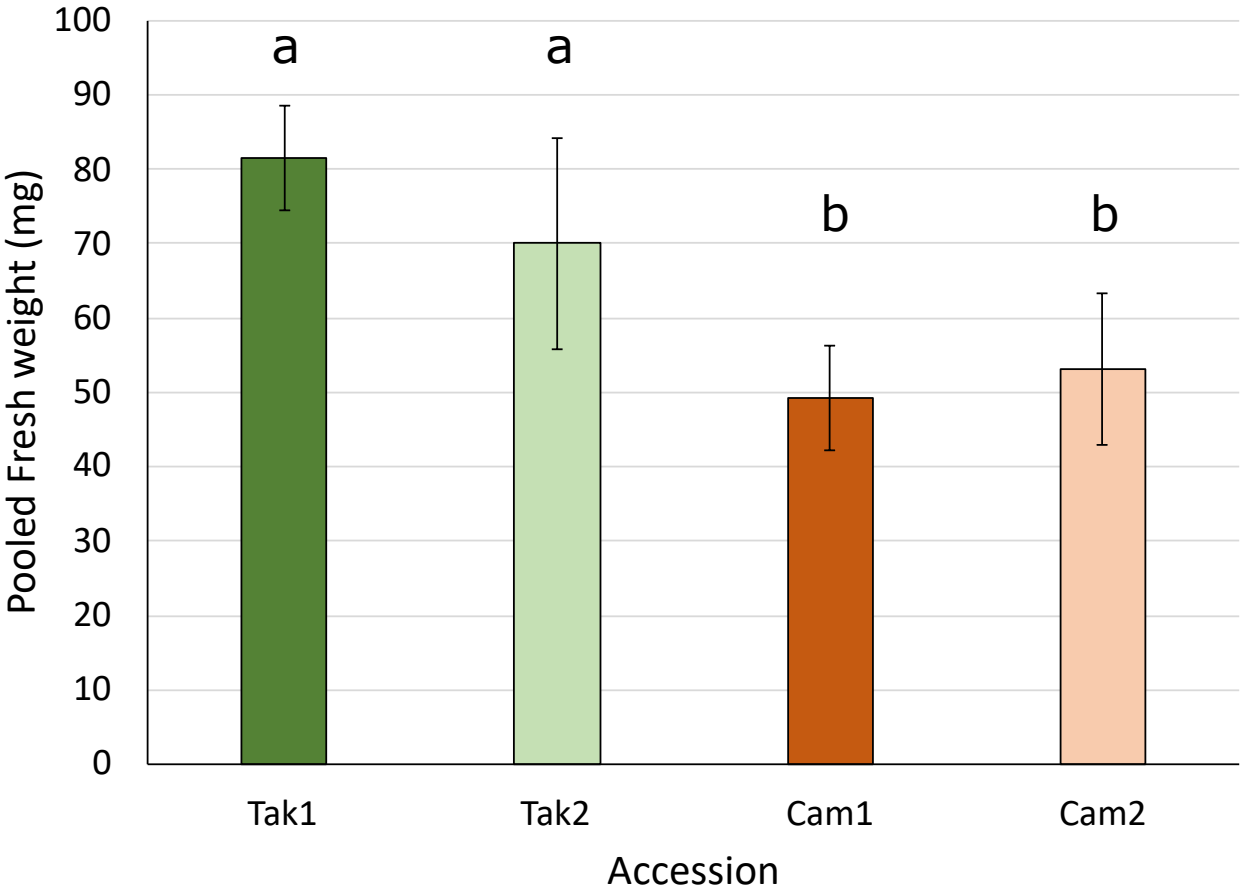


Figure 4.4: Tak1/2 grow faster than Cam1/2 on MS media. 5 Equal sized cuttings of the different accessions (Tak1, Tak2, Cam1, Cam2) of *M.polymorpha* were plated on 1/2 MS and have different growth rates (n=6) based on pooled fresh weight of the 5 cuttings.

To assess if there was a difference in growth between the two wildtype accession of *M. polymorpha* 5 consistent sized cuttings of *M. polymorpha* were used and plated on 1/2 MS and sucrose (1%) for 6 biological replicates with no additional stress and growth was measured after a week and therefore 30 plants cuttings were accessed. Tak1 and Tak2 had no significant difference in terms of growth and therefore there seems to be no difference in terms of fresh weight between the sexes in Tak1/2 plants. Cam1 and Cam2 plants also had no significant difference between the sexes within that accession however there was a significant difference in terms of growth

after a week between Tak1/2 accession and Cam1/2 accession ($P < 0.05$, n=6, 4.4). Therefore assessment of salt growth phenotype was normalised to the 0 mM salt condition on the same accession.

4.2.4 *M. polymorpha* Accessions have a Similar Salt Tolerance

To assess if there is a difference in salt tolerance between the accession when growth differences are normalised for 5 consistent sized cuttings of each accession was plated on 1/2 MS and sucrose (1%) with 0, 50, 100, 150 mM NaCl and pooled fresh weight of the 5 cuttings was accessed (n=6). For analysis each biological replicate was normalised to 0 mM salt. 50, 100, 150 mM salt conditions weighed significantly less than each preceding condition in all accessions other than Tak1 under 50 mM NaCl ($P<0.05$, 4.5). However all accessions had the same reduction in mass when normalised to their own 0 mM salt condition. Therefore all accessions showed the same reduction in growth under salt stress in 50, 100, 150 mM NaCl but there was no significant difference in salt tolerance despite differences in growth between the Cam and Tak accessions under 0 mM salt conditions.

4.2.5 CDPK- E and CDPK-F are the Only CDPKs Upregulated in Response to Salt Stress in Tak1

qRTPCR was used to access the changes in expression of each CDPK in response to salt in adult plants. *CDPK-A/B/C/G* all seem to be significantly downregulated in response to salt at 50 mM NaCl in Tak1 but seem to not change in the 100 or 150 mM NaCl condition. Therefore *CDPK-A/B/C/G* were not found to be modified in a dosage dependent manner and are not considered to be salt responsive. However, it could also be that they are involved in low salt conditions. *CDPK-E* and *CDPK-F* are both upregulated in response to NaCl in the Tak1 accession of *M. polymorpha* in a dosage dependent manner. Expression of *CDPK-E* was upregulated 6-fold in 150

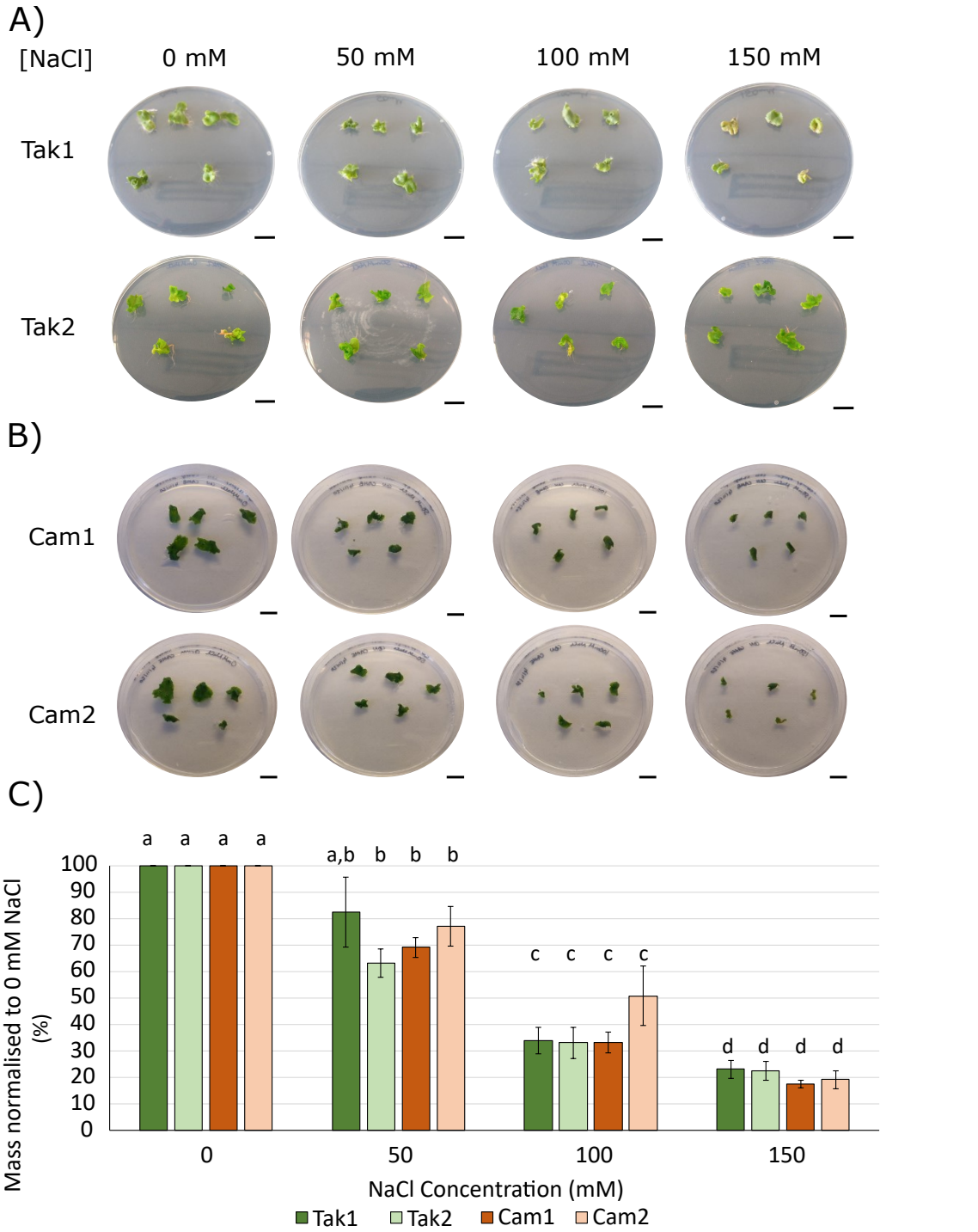


Figure 4.5: Cam1/2 have the same salt tolerance as Tak1/2 *M. polymorpha*. Different Accessions of *M. polymorpha* have a similar salt tolerance based on growth when treated with salt concentrations from 0-150 mM. A) Images of plates containing 5 cuttings of Tak1 and Tak2 accessions grown for 1 week on 0-150 mM NaCl. B)Images of plates containing 5 cuttings of Cam1 and Cam2 accessions grown for 1 week on 0-150 mM NaCl. C) Pooled fresh weight of 5 samples of each accession of *M. polymorpha* (n=6) normalised to 0 mM NaCl to access for growth under 50-150 mM NaCl treatment

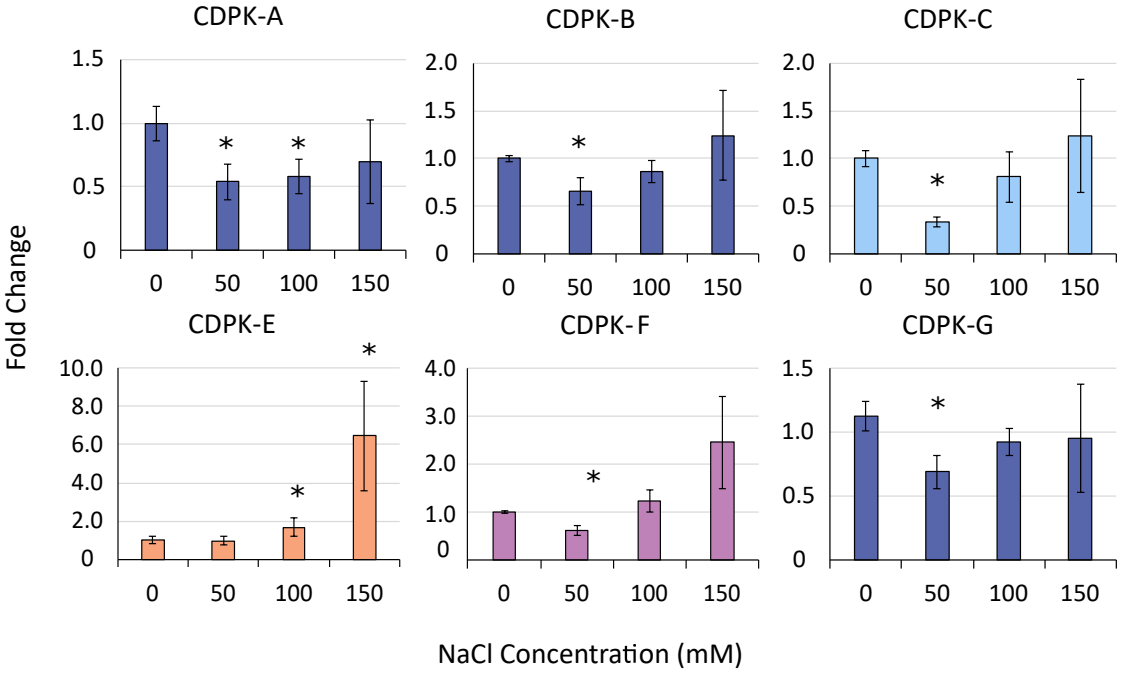


Figure 4.6: Expression Analysis of CDPK-A/B/C/E/F/G in Tak1 *M. polymorpha*. qRT-PCR was used to analyse expression with 3 technical replicates and 3 biological replicates and 2 internal controls (Actin and Adenine phosphoribosyltransferase). Each CDPK is coloured based on subgroup from phylogeny (Fig.3.1). Error bar are +1/-1 SE. *= P<0.05, **=P<0.01.

mM NaCl and also seems to be slightly upregulated at 100 mM condition. CDPK-E is a member of subgroup IV CDPKs and therefore in the same subgroup as AtCPK16 which is involved in the maintainence of salt induced signals through phosphorylation of GLR3.7 (Wang et al. 2019). None of the other CDPKs in subgroup IV are thought to be involved in salt induced signals in *A. thaliana* so CDPK-E may be a homologue of AtCPK16, or CDPK-E may be novel in function. *CDPK-F* is downregulated in 50 mM salt similar to the other CDPKs and is upregulated in 100 mM salt and 150 mM salt to a lesser extent than CDPK-E. Interestingly CDPK-F is encoded in the V chromosome of male Tak1 and CDPK-D was the homologue found on the U chromosome of Tak2 which is why CDPK-D could not be accessed in the Tak1 qRTPCR. Of the *A. thaliana* subgroup 3 CDPKs, only AtCPK10 is involved in salt tolerance through interaction with HSP1 and reduces stomatal conduction to defend against the osmotic stress component of salt stress. As *M. polymorpha* has air pores and not stomata it is difficult to suggest a function based on this as it

is not currently known if air pores can reduce in conduction due to the difficulty of investigating the cells at the base of the air pore.

4.2.6 LEA1 and CIPK-B are Modified in Expression in Tak1 *M. polymorpha* under Salt Stress

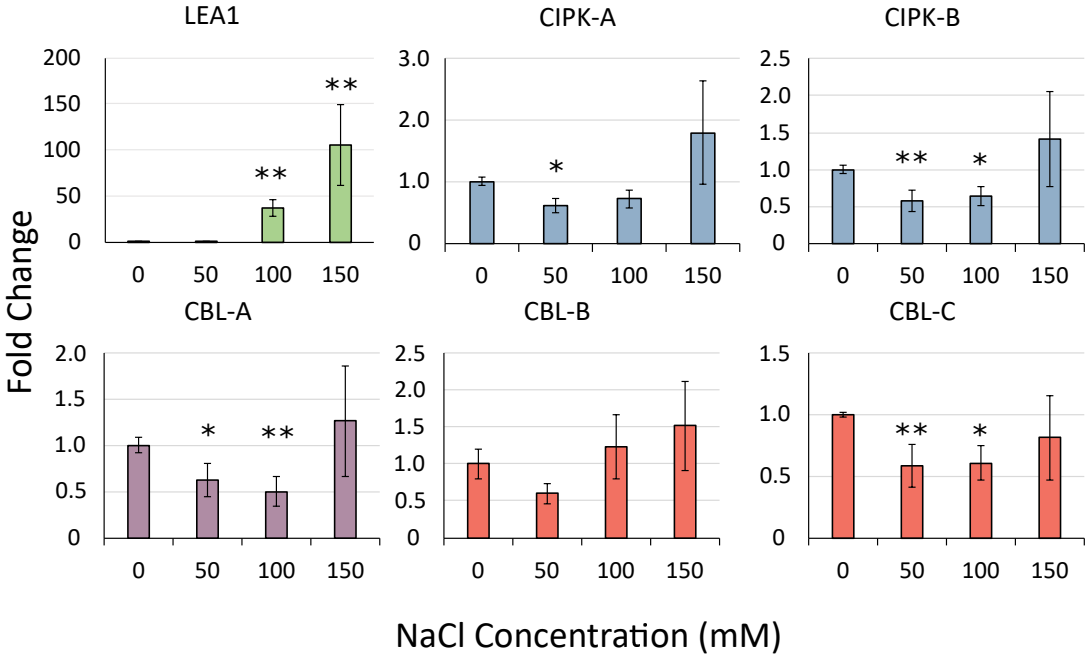


Figure 4.7: Expression Analysis of CBL-A/B/C, CIPK-A/B and LEA1 in Tak1 *M. polymorpha*. Expression was analysed with qRTPCR with 3 technical replicates and 3 biological replicates and 2 internal controls (Actin and Adenine phosphoribosyltransferase). Each gene is coloured based on subgroup from phylogeny (Fig.3.2; Fig.3.4). Error bar are +1/-1 SE. *= P<0.05, **=P<0.01.

LEA1 and *CIPK-B* are modified in expression in Tak1 accession of *M. polymorpha*. qRTPCR was used to access the changes in expression of a pooled sample of 5 plants for 3 technical and 3 biological replicates normalised to actin and ADP. *LEA1* was used to access the effects of osmotic of the salt stress conditions as a positive control and was significantly upregulated at 100 mM (P<0.001) and upregulated 100 fold in 150 mM (P<0.01). *LEA1* did not significantly change in expression at 50 mM NaCl which may be because LEA1 predominantly responds to the osmotic stress which may be not be a factor at 50 mM however the ionic effects of the stress must still be significant as there were significant changes in phenotype under

that condition. *CIPK-A* was only significantly downregulated at 50 mM ($P<0.05$) and therefore it cannot be described as dosage dependent but it could be true that CIPK-A specifically responds under low salt conditions but not high salt conditions or specifically under drought conditions. CIPK-B is a more likely NaCl stress responsive candidate in Tak1 as it is downregulated in 50 mM ($P<0.01$) and 100 mM ($P<0.05$) conditions.

4.2.7 CBL-A and CBL-C are downregulated in Tak1 *M. polymorpha* under Salt Stress

CBL-A and *CBL-C* are both downregulated at 50 and 100 mM NaCl in the Tak1 accession of *M. polymorpha*. qRTPCR was used to assess the changes in expression of a pooled sample of 5 plants for 3 technical and 3 biological replicates normalised to actin and ADP. *CBL-A* is downregulated in 50 mM ($P<0.05$) and 100 mM ($P<0.01$) NaCl and therefore could be involved in NaCl signalling in Tak1. As CBL-A is predicted to localise to the vacuole it may be that it is homologous to AtCBL2/3 and interact with a V-ATPase to control vacuolar pH to help with the sequestration of ionic components of salt stress e.g. Na^+ and Cl^-. Similarly CBL-C is downregulated in 50 mM ($P<0.01$) and 100 mM NaCl ($P<0.05$) and as it is predicted to be plasma membrane localised it is a likely AtSOS3 (AtCBL4) homologue which may suggest that it interacts with a Na^+/H^+ antiporter similar to the SOS pathway in *A. thaliana*. Interestingly none of the CBLs or CIPKs were significantly downregulated at 150 mM salt conditions which may suggest that lower salt conditions have a different mechanism of salt tolerance than high salinity conditions. This may be further evidenced by the fact that *LEA1, CDPK-E* and *CDPK-F* are not significantly modified in expression in 50 mM but are at higher concentrations of salt in Tak1.

4.2.8 LEA1, CIPK-A and CIPK-B are Modified in Expression under Salt Stress in the Cam2 Accession

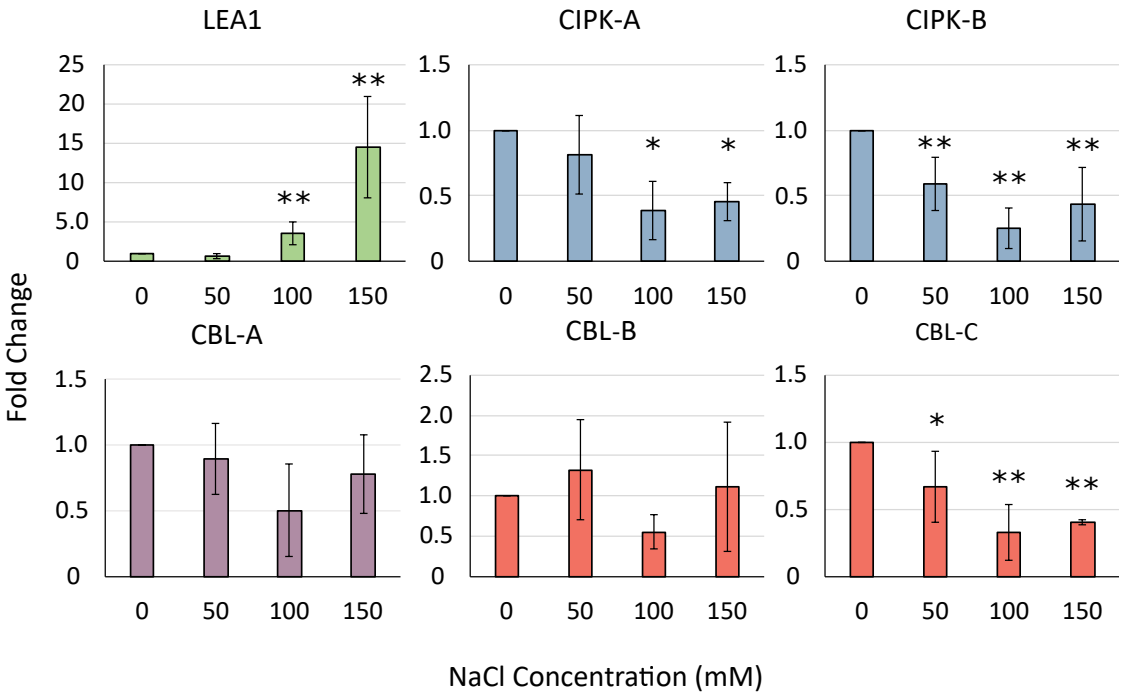


Figure 4.8: Expression Analysis of CBL-A/B/C, CIPK-A/B and LEA1 in Cam2 *M. polymorpha*. Expression analysis with qRTPCR with 3 technical replicates and 3 biological replicates and 2 internal controls (Actin and Adenine phosphoribosyltransferase). Each gene is coloured based on subgroup from phylogeny (Fig.3.2; Fig.3.4). Error bar are +1/-1 SE. *= P<0.05, **=P<0.01.

The Cam2 accession has modified expression for both CIPKs in *M. polymorpha* as well as *LEA1* in a dosage dependent manner under salt stress. *LEA1* is a osmotic stress gene upregulated under salt stress as well. In Cam2 *LEA1* is upregulated 4 fold in plants treated with 100 mM (P<0.01) and 15 fold in plants treated with150 mM (P<0.001) NaCl stress for 1 week therefore it is likely that the plants are under the osmotic effects of salt stress from 100 mM NaCl. *CIPK-A* is downregulated under 100 mM (P<0.05) and 150 mM (P<0.05) which may suggest that it only involved in high salinity stress or that it is involved in osmotic stress components of salt stress. *CIPK-B* is downregulated in a dosage dependent manner in Cam2 in 50, 100 and 150 mM NaCl (P<0.001) which may suggest its involvement in salt stress even in relatively low salt stress conditions (50 mM). Therefore based on this qRT-PCR result it is likely that CIPK-B is involved in salinity stress through

ionic stress involvement. However, it is also possible that *CIPK-A* is involved as it is downregulated in 100 mM ($P<0.05$) and 150 mM salt ($P<0.05$) but further experimentation would be required to know the exact involvement of the CIPKs in *M. polymorpha* in salt stress.

4.2.9 CBL-C is the only CBL downregulated in Cam2 *M. polymorpha*

Neither *CBL-A* or *CBL-B* show altered expression in Cam2 plants treated with salt for 1 week. *CBL-C* is the only CBL modified in expression in Cam2 and is downregulated under 50 mM ($P<0.05$), 100 mM ($P<0.001$) and 150 mM ($P<0.001$) NaCl. This CBL is predicted to be localised to the plasma membrane and as *CBL-C* is modified at 50 mM salt, which *LEA1* was not modified under, it seems that *CBL-C* is likely responding to the ionic effects of salt stress which may be either Na^+ or Cl^-. As previously described the SOS3 (AtCBL4) interacts with an Na^+/H^+ antiporter and therefore it may be that CBL-C is a SOS3 homologue in *M. polymorpha*.

4.2.10 A Public RNA Sequencing Dataset confirms CBL-C and CDPK-E, but not the CIPKs, as Salt Regulated after 1 week treatment

A publically available dataset of RNA sequencing was available for Tak1 *M. polymorpha* two week old immature thalli grown on Gamborg's B5 medium (1/2 strength) with 50 mM NaCl treatment (Tanaka et al. 2018). Originally this dataset was analysed using version 3.1 of the *M. polymorpha* genome so it was reanalysed with the version 5.1 and analysed for potential changes in expression in CBLs, CIPKs and CDPKs (Salmon and DESeq2; Chapter 2.7.1). From this analysis, *CBL-A* and *CBL-B* did not seem to be downregulated but may actually be upregulated slightly. *CBL-C* is still downregulated under these conditions which is also true for the Tak1 and Cam2 qRTPCR and therefore it is a very good candidate for a salt

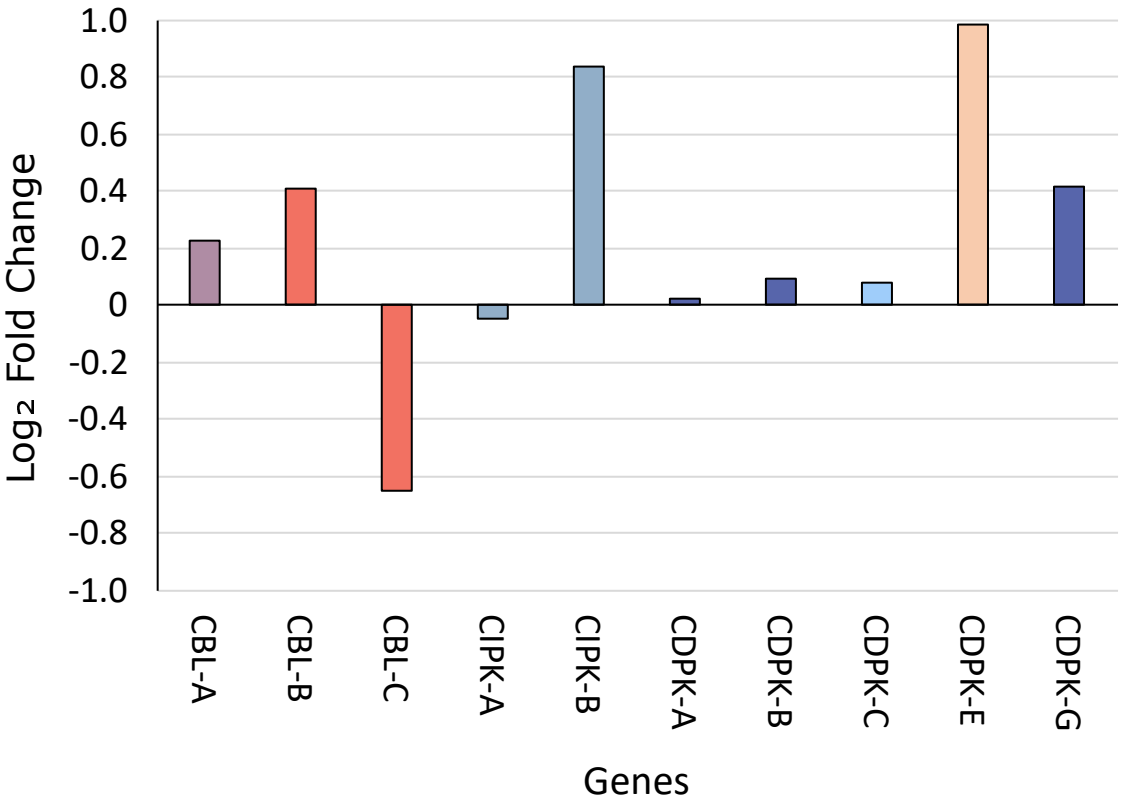


Figure 4.9: RNA sequencing analysis of *M. polymorpha* on 50 mM Salt for 1 week. RNA Sequencing Analysis using Salmon and DESeq2 against *M. polymorpha* v5.1 transcriptome for MpCBLs, MpCIPKs, MpCDPKs from Tanaka et al., (2018) dataset after 1 week 50 mM NaCl Tak1 plants on 1/2 Gamborg Media.

responsive CBL. *CIPK-A* did not show any large change in expression under this condition and *CIPK-B* seemed to be upregulated in Tanaka et al. (2018) unlike the qRT-PCR which demonstrated downregulation in mature thalli in CIPK-B. Of the CDPKs, *CDPK-E* is upregulated in this dataset and *CDPK-G* may also be slightly upregulated.

4.2.11 A Public RNA Sequencing Demonstrates a Different Early Salt Response in Tak1 Calcium Decoders

Another publically available RNA sequencing dataset has been released for 0, 2 and 24 hour salt responsive transcriptional change analysis in Tak1 *M. polymorpha* gemmaeling grown on Gamborg's B5 medium (1/2 strength) for 1 week and treated with 100 mM NaCl (Wu et al. 2021). For gene expression to be associated with salt tolerance it would be expected to have a consistent increase in expression upon

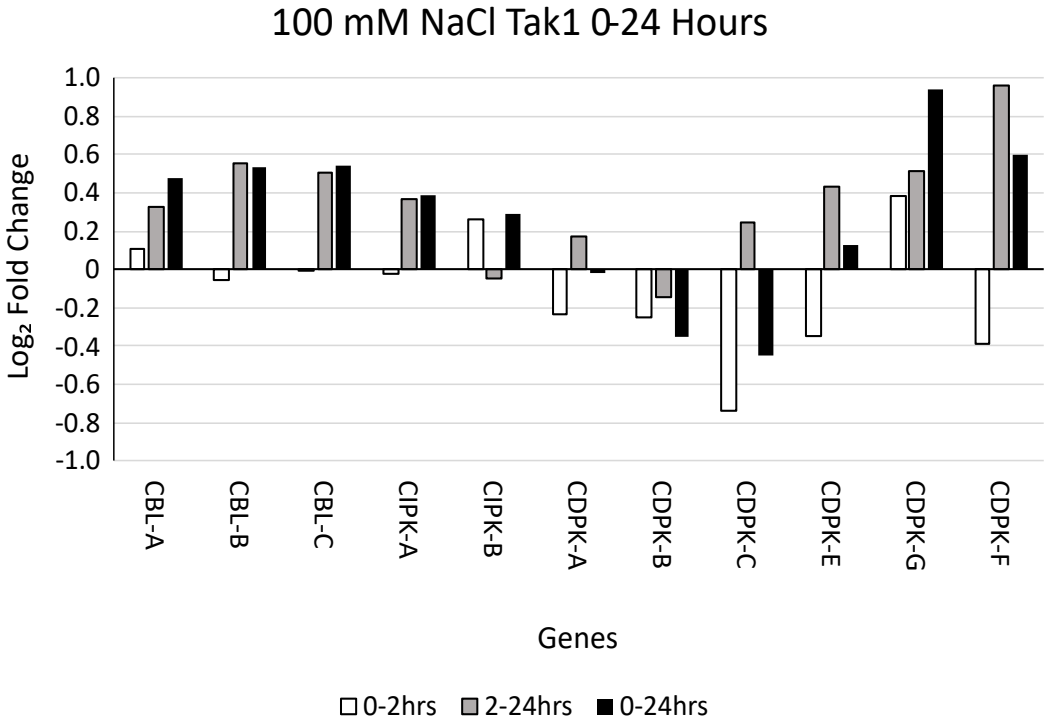


Figure 4.10: RNA Sequencing Analysis using Salmon and DESeq2 against *M. polymorpha* v5.1 transcriptome for MpCBLs, MpCIPKs, MpCDPKs from Wu et al., (2021) dataset including 0-2 hours (White), 2-24 hours (Grey) and 0-24 hours (Black) 100 mM NaCl Tak1 plants on 1/2 Gamborg Media.

a salt stimulus. Not all genes will modify in expression immediately as some will be later responders to salt stress. All the *CBLs (A/B/C)* are upregulated from 0-24 hours under salt stress. *CIPK-A* is upregulated from 2-24 hours and therefore seems to be linked to early salt response. *CIPK-B* seems to be upregulated from 0-2 hours and downregulated from 2-24hrs and therefore is likely not an early responder to salt stress. *CDPK-C/E/F* are all changing between down regulation from 0-2 hours and then upregulated from 2-24 hours and the inconsistency between these responses makes them unlikely to be early salt responders. *CDPK-B* is consistently downregulated in early salt response which is unexpected, as typically this would make it likely a negative regulator of salt tolerance being repressed under salt stress, however other CDPKs that are negative regulators in *A. thaliana* are upregulated initially and then subsequently downregulated post 24 hours. *CDPK-G* is upregulated consistently from 0-2 then 2-24 hours which suggests it may have some role in early plant salt tolerance responses.

4.3 Discussion

4.3.1 Phenotypic comparison of M. polymorpha accessions under abiotic stress

Calcium decoders respond to salt stress to mitigate the potential ionic and osmotic components of the stress in plants such as *A. thaliana* and *O. sativa*. CBL and CIPK pathways deal with ionic stress responses predominantly through control of ion channels such as with the Salt Overly Sensitive pathway which operates on the SOS1 Na^+/H^+ antiporter to exclude Na^+ from the plant in response to salt stress. Other CBLs and CIPKs are involved in counterbalancing ions such as K^+ or controlling vacuolar pH to allow for sequestration of potentially damagining ions in the vacuole. CDPKs are known to control osmotic stress responses in plants through interactions with ABA, ROS production, and through controlling stomatal conduction. Therefore it was likely that some of the calcium decoders involved in *M. polymorpha* would also have involvment in salt stress responses. Potentially some of these calcium decoders may work through conserved pathways found in other plants which will garner insight in the evolution of plant tolerance conferred by calcium decoders however some may be novel to *M. polymorpha* which would provide biotechnological options for creating new salt tolerance pathways in crop plants.

In order to determine salt stress tolerance in *M. polymorpha* the tolerance of *M. polymorpha* to salt first had to be established. Initially the tolerance was assessed through placing cuttings of adult *M. polymorpha* on salt concentrations from 0-200 mM and it was found that increased salt resulted in chlorosis of adult plants and some production of anthocyanins at low concentrations of salt but adult plants survived concentrations of 0-150 mM NaCl in Tak1 (Fig.4.2). Studies by Tanaka et al. (2018) confirmed that Tak1 survives 0-100 mM but does not survive 200 mM after 1 week, but 150 mM NaCl was not assessed. To get more consistent plant sizes and to ensure that plants were in the same life phase Tak1 plants gemmae were

grown for 1 week then exposed to salt stress for 1 week. It was found that plants could survive up to 200 mM NaCl still survive with some green tissue which may be due to differences in life phase as previously studies have shown that gammae retain some dormancy responses presumably to increase dispersal chances (Eklund et al. (2018); Fig.4.3). It was also confirmed that, similar to other plants such as *A. thaliana*, Tak1 is less sensitive to sorbitol caused osmotic stress than salt stress from either KCl or NaCl. Sorbitol stressed plants demonstrated less chlorosis at 400 mM than salt stressed plants at 200 mM despite equally reduced water potential (Fig.4.3). Therefore the ionic effects of the salt stress demonstrably cause the lack of survivability beyond 150 mM salt and not just osmotic effects of salt stress.

Tak1 *M. polymorpha* has distinct calcium signals in response to salt and sorbitol. Previous work in the laboratory has transformed Tak1 *M. polymorpha* with an R-GECO construct as a fluorescent reporter to detect calcium signals at a single cell resolution (Fig.4.1). The R-GECO line demonstrated a biphasic peak in response to 100 mM NaCl stress that was not detectable when calcium ion channels were blocked by lanthanum chloride, and a single peak was detected in response to 200 mM sorbitol with both calcium signals taking place over 40 minutes (Fig.4.1). The calcium signal detected in *M. polymorpha* was similar to those found in *A. thaliana* in terms of a single peak to sorbitol and a biphasic peak in response to salt stress however the signal in *A. thaliana* took place over only 3 minutes (Schmöckel et al. 2015). Studies in *A. thaliana* have demonstrated that the calcium signal in response to salt takes place in the nucleus and the cytoplasm and that these signals are independent of each other as blocking the signal in one location does not block the signal taking place at the other intracellular location. Despite the similar time-frame of both osmotic and salt induced signals: blocking the signal in response to osmotic stress did not stop root growth inhibition, but blocking both phases of the signal in response to salt did stop root growth inhibition (Huang et al. 2017). Therefore it is likely that the decoding components activated by the calcium signal are distinct between osmotic stress by itself, and salt stress induced osmotic and

ionic stress in *A. thaliana*. Hence we would not neccesarily expect that all of the salt stress tolerance calcium decoders would be activated under an osmotic stress induced calcium signal.

Multiple accessions of *M. polymorpha* are available for study including Tak1/2 from Japan and Cam1/2 from the UK, and due to different sources, may have different levels of salt tolerance. Therefore, it may be that there are differences in calcium signal induced tolerance responses which can be investigated. Similarly some studies have shown different osmotic stress tolerance in bryophytes dependent on sex of the plant and therefore it was important to assess if there is a difference in terms of growth and salt tolerance dependent on accession and sex of the *M. polymorpha* (Marks et al. 2016). Taking cuttings of adult tissue of similar sizes for Cam1/2 and Tak1/2 demonstrated no difference in growth between the sexes when not under stress conditions but a significantly different fresh weight between Cam and Tak accessions (Fig.4.4). Therefore to assess differences in stress tolerance between the accessions, fresh weight cannot be used and each replicate was normalised to an unstressed condition. When plants were stressed with 50-150 mM NaCl for one week no difference was found between accessions in terms of fresh weight and therefore they have a similar salt tolerance between Cam and Tak but different rates of growth generally. Hence both accession and sexes of *M. polymorpha* can survive conditions of 0-150 mM NaCl with a similar tolerance.

4.3.2 Transcriptional Changes of CDPKs in *M. polymorpha* under Salt Stress

Transcriptional changes of CDPKs in response to salt stress vary in terms of initial stress response (0-32 h) and continued stress over a prolonged time (post day 6). Under 100 mM NaCl stress Tak1 plants downregulate *CDPK-A/B/C/E/F* from 0 to 2 hours which would not imply function in salt tolerance responses if compared to expression changes in *A. thaliana* whereby positive and negative regulators of salt stress tolerance are upregulated initially in reponse to salt stress. However,

considering that the salt induced calcium signals takes 40 minutes in *M. polymorpha*, unlike *A. thaliana* which has a 3 minute signal, it may be that the responses of *M. polymorpha* to salt are slower and therefore the transcriptional changes may be more representative from 2-24 hours. From 2-24 hours *CDPK-C/E/F/G* are all upregulated and, of these, *CDPK-G/F* are upregulated greater than 0.6 log2 fold change. CDPK-F is encoded on the V chromosome and a member of subgroup III and therefore, of the known AtCPKs, it is most similar to AtCPK10 which interacts with HSP1 to control stomatal conductance. However, it is not known that air conduction is controlled in air pores found in *M. polymorpha* therefore CDPK-F may have a novel function in early salt response (Zou et al. 2010). However after 1 week of 50 mM NaCl treatment CDPK-F could not be detected to have a significant change in response to salt (Fig.4.6) but qRT-PCR confirmation demonstrates that it is upregulated at higher salt concentrations (Fig.4.6). CDPK-G is a member of the subgroup I CDPKs and a number of the AtCPKs from this subgroup are involved in regulating ABA based drought responses such as AtCPK11/4/6 which are also involved in osmotic components of stress. The 50 mM NaCl treatment by Tanaka et al. (2018) also shows upregulation of *CDPK-G* after 1 week which makes it very likely to be a salt responsive gene but this change in regulation could not be confirmed through qRT-PCR (Fig.4.6).

Of the CDPKs found in *M. polymorpha CDPK-E* is the best candidate as a salt responsive gene likely to mediate salt tolerance. CDPK-E is a member of subgroup IV and has most similarity to AtCPK16 which is involved in the continuation of calcium signals in response to salt phosphorylation of GLR3.7 channel to mediate calcium release (Wang et al. 2019). *CDPK-E* is also consistently upregulated from 2-24 hours (Fig.4.10) at 100 mM NaCl, 1 week post treatment with 50 mM NaCl (Fig.4.9) and with qRT-PCR 100-150 mM NaCl (Fig.4.6). In contrast to this CDPK-G could also be a interesting target as a member of subgroup I and upregulated in all RNA sequencing datasets but it could not be confirmed using qRT-PCR which may imply either heterogeneity between Tak1 in the three labratories or the change

in expression may be reduced due to the different medias used for the experiments as both RNA sequencing datasets used Gamborg media but the qRT-PCR experiment was carried out on plants grown on MS media. CDPK-F could also be an interesting target as it is located on the V chromosome of Tak1 and is homologous to CDPK-D found in Tak2. However no difference was between Tak1 and Tak2 in terms of salt tolerance from fresh growth measurement (Fig.4.4). Similarly while *CDPK-F* could be detected by qRT-PCR changes in expression (Fig.4.6), and from 2-24 hours with RNA sequencing (Fig.4.10), it was not detectable in the Tanaka et al. (2018) dataset as a significantly changed gene after 1 week 50 mM treatment. *CDPK-F* was also downregulated from 0-2 hours treatment significantly (Fig.4.10). This may also mean that there is heterogenity between Tak1 accessions or that the modification of *CDPK-F* expression may be different at different life phases. The Tanaka et al. (2018) experiment was similar to the qRT-PCR experiment in design but grew plants on Gamborg instead of MS media, and used immature plants ($<$ 3 weeks old) which may have caused the difference in *CDPK-F* expression. The Tanaka et al. (2018) experiment was also very similar to Wu et al. (2021) and used the same media and plants of the same age. However, due to length of the treatment the plants that were accessed for changes in expression were 6 days older in a 3 week lifecycle to become adult plants. The change in ages of the plant may have affected the changes in expression of *CDPK-F* if the gene is involved in development. It could also be that *CDPK-F* changes in gene expression predominantly take place in response to higher salt shock as both (Wu et al. 2021) and the the qRT-PCR experiment detected changes at 100 mM NaCl whereas Tanaka et al. (2018) only used a 50 mM NaCl treatment. Therefore the most consistent changes in expression took place in *CDPK-E* in all treatments lengths, treatment strengths, media and lifephase.

4.3.3 Transcriptional Changes of CBLs in *M. polymorpha* under Salt Stress

CBL-C is the most likely salt sensitive CBL candidate in *M. polymorpha*. CBL-A was unlikely to directly act on salt induced calcium signals due to its predicted localisation (Chapter 3.2.2) to the tonoplast whereas most salt induced signals in *M. polymorpha* are thought to take place at the plasma membrane and potentially the nucleus as the detected R-GECO signal was mostly detectable at the plasma membrane and signals in *A. thaliana* take place at the PM and nucleus (Fig.4.1). However CBL-A is similar to AtCBL2/3 which controls a V-ATPase to moderate the pH of the vacuole which may have important functions in salt tolerance to help sequester Na^+. Also *CBL-A* is upregulated from 0-24 hours (Fig.4.10) and slightly upregulated in RNA Sequencing after 1 week (Fig.4.9) but was not modified in Cam2 and was downregulated in Tak1 qRT-PCR datasets (Fig.4.7; Fig.4.8). CBL-B is predicted to localise to the plasma membrane and is upregulated from 2-24hours (Fig.4.10), and after 1 week by RNA sequencing (Fig.4.9) but this result could not be confirmed by qRT-PCR in either Tak1 or Cam2 (Fig.4.7; Fig.4.8) which implies it is unlikely to be involved in salt induced calcium signalling responses. *CBL-C* is upregulated from 2-24 hours (Fig.4.10), and downregulated after 1 week RNA sequencing (Fig.4.9). *CBL-C* is also downregulated consistently in both Cam2 and Tak1 after 1 week salt treatment from 50-150 mM in a dosage dependent manner which acts as strong evidence that it is salt dependent regardless of media, lifephase, and accession of *M. polymorpha*.

Previous studies in *A. thaliana* have shown that *Salt Overly Sensitive 3*, a salt responsive CBL, is upregulated in specific cell subtypes in the root in the first 24 hours of treatment (Ji et al. 2013) but are downregulated at 6 days (Rolly et al. 2020). As *CBL-C* is upregulated from 2-24 hours in *M. polymorpha* but downregulated at day 7 it is possible that CBL-C is a homologue of AtSOS3. This would also be interesting as it would mean that SOS3 expression regulation is conserved from *M. polymorpha* to *A. thaliana* which are seperated evolutionary paths over 400 million years

ago. Similarly *CBL-C* seems to be transcriptionally regulated the same way in both accessions from the UK and Japan which demonstrates a very wide environmental range that the plants share a conserved mechanism. The change in regulation is also interesting in itself as it may demonstrate that CBL-CIPK regulation of Na^+/H^+ antiporter AtSOS1 is an early salt tolerance response to increase activity of the ion channel which is downregulated under continued stimulus. However more than just similar transcriptional response is required to confirm the hypothesis that CBL-C is a SOS3 homologue.

4.3.4 Transcriptional Changes of CIPKs in *M. polymorpha* under Salt Stress

It is difficult to propose a MpCIPK that is salt responsive based on transcriptional information in *M. polymorpha*. Based on the qRT-PCR experiments in Tak1 *CIPK-B* seems to downregulated at 50 and 100 mM NaCl and in Cam2 CIPK-B is downregulated from 50-150 mM after 1 week which may imply that it is a salt responsive CIPK in *M. polymorpha* (Fig.4.7; Fig.4.8). However the RNA sequencing information shows downregulation of *CIPK-B* in Tak1 from 0-24 hours but it seems to be upregulated after 1 week under 50 mM stress (Fig.4.10; Fig.4.9). The RNA sequencing dataset that shows upregulation differs in media that the plants are grown on and lifephase as mentioned before which may imply that these transcriptional difference may be due to these differences. Similarly the lack of downregulation at 150 mM in Tak1 qRT-PCR shows a lack of dosage dependence in this dataset for this gene which makes it difficult to suggest that this gene is salt responsive. *CIPK-A* is upregulated from 2-24 hours but doesnt show much change after 1 week of 50 mM NaCl in terms of expression (Fig.4.10; Fig.4.9). Based on qRT-PCR data *CIPK-A* is slightly downregulated at 50 mM but not at higher concentrations of salt and in Cam2 *CIPK-A* is downregulated at 100-150 mM NaCl (Fig.4.7; Fig.4.8). The lack of large and consistent change also makes it difficult to ascribe CIPK-A as a salt responsive CIPK.

CIPKs may not change much in terms of expression in response to an activating salt stimulus as a CBL is required for activity. The CIPKs in *M. polymorpha* did not demonstrate large consistent changes in response to salt in a dosage dependent manner. However, considering that the *CBL-C* expression is downregulate in response to salt stimulus may be sufficient to reduce activation of the a salt responsive *M. polymorpha* CIPK and to stop induction of activity in downstream ion channels. This model would imply that in response to salt a calcium signal is generated which activates a CBL, and potentially CDPKs. The CBL then interacts with a CIPK to modulate its activity initially to control plant salt tolerance responses and this is reinforced by increased expression of the activating CBL. Post 24 hours expression of the CBL is reduced as other defence responses take over and the reduced number of CBLs inhibits the activation of the CIPKs to continued salt induced calcium signals. Therefore changes in expression of the CIPK is not required as the pathways is controlled at the calcium sensor level. As some CIPKs are known to be activated by other CBLs to have involvement in multiple pathways, this control of the pathway at the CBL level may mean that a sufficient amount CIPK protein is still present to be involved in other signalling pathways that it may be involved in.

Chapter 5

Comparison of Interactions between CBLs and CIPKs

5.1 Deconvoluting the *M. polymorpha* CBL-CIPK network through Interaction Testing

Calcineurin B-Like proteins were originally described in plants and thought to be unique calcium sensors that interact with protein kinases known as CIPKs. Calcineurin B-Like proteins contain 4 calcium binding domains similar to calmodulin but crystal structures have demonstrated a similar folding structure to calcineurin B proteins found in animals and fungi (Negae et al. 2003; Kumar et al. 2016). Calcineurin B proteins physically interact with Calcineurin A proteins which also interact with calmodulin to activate the Calcineurin A phosphatase activity. However Calcineurin A phosphatases are not present in plants which were thought to have either coevolved a similar structure CBL protein or lost Calcineurin A and evolved other interactors (Shi et al. 1999). The first study to find an interactor for a plant CBL used a library of proteins from *A. thaliana* and found a specific kinase they called CBL-Interacting Protein Kinase (CIPK) through yeast two-hybrid and *in vitro* binding assays (Halfter et al. 2000). Futher studies have since found that CBLs and CIPKs are also present in other bikonta including archaeplastida,

chromalveolates and excavates, and another CBL-CIPK complex has been functionally characterised in the human pathogen *Trichomonas vaginalis* (Beckmann et al. 2016). Therefore CBLs and CIPKs, while initially described in plants and thought to be a unique evolution from Calcineurin B it is present in other bikonta species as well.

5.1.1 CIPKs interact with CBLs through the NAF domain

The calcium signal that stimulates a CBL is transduced through to the CIPK through a physical interaction and activates the kinase activity of the CIPK to phosphorylate downstream protein. The physical nature of this interaction was confirmed with both the yeast two-hybrid and affinity chromatography of the AtCBL1 protein using *A. thaliana* protein extract (Shi et al. 1999). Subsequent studies determined that a novel interaction domain, the 'NAF domain', of the CIPK was necessary and sufficient to mediate the interaction with a CBL. The NAF domain, named for its most highly conserved residues, was identified using Yeast two-hybrid with reduced sections of the C-terminal section of the CIPK identified by Shi et al. (1999) until it was defined as 24 amino acids (Albrecht et al. 2001). Mutations of the N- and C- terminal parts of the NAF domain abolished the interactions with the CBL but mutation in the centre of the domain only reduced interaction (Albrecht et al. 2001). Crystal structures have since been made of the C-terminus of SOS2 and SOS3 which demonstrated that the NAF domain forms hydrophobic interaction between 12 amino acids in SOS3 and 7 amino acids of SOS2 which is stabilised by 5 hydrogen bonds. The interaction with this domain forces SOS2 into an active confirmation (Sánchez-Barrena et al. 2007).

The interaction between CBL and CIPK is specific and confers localisation to allow for kinase activity targetted to downstream proteins. The SOS3(G2A) mutation blocks N-myristoylation which means that SOS3 can no longer localise to the plasma membrane and overexpression of this mutant in the *sos3-1* background does not complement the salt sensitivity phenotype. The lack of function of SOS3(G2A)

is likely due to a lack of localisation as SOS2 contains no intrinsic localisation domains (Ishitani et al. 2000). Similarly AtCBL10 functions similarly to SOS3 however it is expressed in the shoot instead of the root but it is typically localised to the tonoplast. Upon stimulation AtCBL10 localises to the plasma membrane and interacts with SOS2 in order to activate SOS1. Therefore the recruitment to the plasma membrane is important for SOS2 to phosphorylate SOS1 upon a salt induced calcium signal (Quan et al. 2007). Similarly AtCBL10 is phosphorylated by SOS2 which stabilises the interaction between the CBL and CIPK to increase the activity of SOS2. This phosphorylation event may also cause the change in localisation for AtCBL10 as phosphorylated AtCBL10 was only detected in the membrane protein fraction and not the soluble fraction, but further studies are required to conclusively prove this (Lin et al. 2009). Both phosphorylation and localisation are required for proper function of SOS2 with AtCBL10.

5.1.2 CBLs interact with CIPKs through the Calcium binding domain

While the exact interaction for each CBL and CIPK has not been resolved, the structure of the SOS2 NAF domain and SOS3-Ca^{2+} bound form has been demonstrated and may offer some indication of general interactions between CBLs and CIPKs. The NAF domains have two α-helices with a stabilised loop between the helices. These N-terminus helix of the NAF domain forms hydrophobic interactions in between EF hands EF3 and EF4 and the C-terminus helix forms hydrophobic interactions with the EF hands of EF1 and EF2. N-terminus of the loop forms interactions between E4 and F4 helices and the C-terminus of the domain interacts between helices E1 and F1 (Sánchez-Barrena et al. 2007). This crystal structure could not demonstrate the how the differences between occupied and unoccupied SOS3 affect the structure and the interaction between SOS3 and SOS2 but it has been demonstrated that this calcium occupancy affects the kinase activity of SOS2 towards SOS1 (Sánchez-Barrena et al. 2005). However a crystal structure of AtCBL2

demonstrates a different position of EF hands relative to each other, hence the position of EF hands may affect CIPK specificity found between particular CBL-CIPK interactions (Sánchez-Barrena et al. 2005). Both phosphorylation and calcium occupancy could change the structure of the CBL locally as binding of calcium causes changes in the angle of the helices with respect to each other (Réty et al. 2000) and phosphorylation can cause large changes in structure through changes in charge. These changes may be significant enough to allow interactions or reinforce weak interactions as has been described for AtCBL10 interaction with SOS2 but more crystal structures would be required to elucidate this.

Calcium is required for activition of the CBL-CIPK pathway but phosphorylation is only required for certain CBL-CIPK pairs suggesting more than one regulation method. AtCBL4 (SOS3) also interacts with AtCIPK6 to mediate plasma membrane targeting of AKT2 but mutations in EF4 of the CBL block the binding of calcium to the CBL which sebsequently blocks the function of this pathway (Held et al. 2011). Similarly AtCBL4 (SOS3) interaction with AtCIPK24 (SOS2) requires calcium binding and the same EF hand point mutation blocks salt induced activation of the SOS1 ion channel (Qiu et al. 2002). Other CBLs such as AtCBL1/9 also require calcium to activate AtCIPK23 which phosphorylates AKT1 to enhance K^+ uptake (Li et al. 2006). Ca^{2+} occupancy may affect interactions between CBLs and the NAF domain of the CIPK either to weaken or strengthen certain interactions. However calcium occupancy may also cause changes in the kinase domain to modulate its activity as the fourth EF hand of the CBL is predicted to be in close proximity to the CBL-CIPK interaction. Phosphorylation, on the other hand, only seems to be required for certain pairs of CBLs and CIPKs to modulate activity. Previous studies have demonstrated that SOS3 is not phosphorylated by SOS2 for proper function but AtCBL10 is phosphorylated to regulate activity (Lin et al. 2009). However, other studies have demonstrated that SOS2 can phosphorylate AtCBL1, AtCBL4 (SOS3), AtCBL10 but not AtCBL7 (Hashimoto et al. 2012) which seems to contrast the result from Lin et al. (2009) despite both reporting *in*

vitro radiolabelling experiments. AtCBL1 has also been shown to be phosphorylated by AtCIPK1 and AtCIPK23 (Hashimoto et al. 2012). However it is difficult to know whether these phosphorylations represent true *in vivo* phosphorylations due to conflicts in the reports. The phosphorylation of AtCBL1 by AtCIPK23 has been confirmed to affect activity of AtCIPK23 towards AKT1 in *Xenopus* oocytes and S201A and S201D mutations disrupts this activity (Hashimoto et al. 2012). Therefore it can be said that phosphorylation does affect activity towards downstream proteins but it cannot be said how much is based on protein-protein interaction yet.

The majority of known interactions with CBLs are to CIPKs and this has meant that the majority of studies for interactions focus on CBL-CIPK interactions but recent studies have demonstrated that this does not encompass all of the potential interactions and functions for CBLs. Interestingly three CBLs have been demonstrated to interact with proteins other than CIPKs and demonstrate inhibitory activity. AtCBL3 demonstrated inhibitory activity towards a 5'-methylthioadenosine nucleosidase (AtMTAN) but only interacts in the presence of calcium. The interaction between AtMTAN required the full length AtMTAN but only amino acids from 109 to 199 from the AtCBL3 which constrast the full length CBL required for CIPK interaction which implies a different method of interaction (Oh et al. 2008). AtCBL10 and AtCBL7 have also been shown to interact with Translocon of the Outer membrane of the Chloroplasts (TOC34) with only 23 amino acids required of TOC34 for interactions with AtCBL10. AtCBL10 is localised to the chloroplasts for this interaction to take place *in vivo* and the interaction does not require calcium, unlike the interaction between AtCBL3 and AtMTAN, but calcium is required for AtCBL10 to inhibit the GTPase activity of TOC34 (Cho et al. 2016). The interactions desccribed between the CBLs and other proteins all require calcium either for activity or for the interaction and demonstrates that while CBL-CIPK interactions are important for understanding the pathway it does not necessarily capture all of the potential interactions that may take place *in vivo* and therefore CBL-CIPK interactions do not neccesarily capture all of the potential functions of CBLs *in vivo*

either.

5.1.3 Determining the interactions of CBLs to CIPKs in *M. polymorpha*

With 3 CBLs and 2 CIPKs in *M. polymorpha* it is important to determine if all of the CBLs and CIPKs can interact or if there is CBL-CIPK specificity. Based on expression CBL-C is the likely salt responsive CBL in *M.polymorpha* and therefore determining the CIPK it interacts with may help to determine if a CBL-CIPK pathway is involved in salt tolerance similar to the SOS pathway. Similarly determining the interactions of the other CBLs and CIPKs will help to elucidate the number of responses to which this pathway may be able to respond. These interactions can be determined through yeast two-hybrid as well as bimolecular fluorescence complementation in tobacco. It may also be important to determine if the localisation predicted through phylogeny is correct which can be assessed through localisations of biomolecular fluorescence complementation constructs in tobacco. SOS2 (AtCIPK24) and SOS3 (AtCBL4) will also be included in interaction tests to determine if the salt sensitive CBL-C can interact with SOS2 (AtCIPK24). Whichever CIPK interacts with CBL-C may also be able to interact with SOS3 (AtCBL4). This will be particularly helpful in the case that CBL-C can interact with both CIPKs and only one of those CIPKs can interact with SOS3 (AtCBL4). The interaction with SOS may also imply that the SOS pathway is a conserved mechanism from *M. polymorpha* which diverged 450 million years ago from the rest of the plant lineage including *A. thaliana*.

5.2 Results

5.2.1 Yeast Two Hybrid determined that CBL-C and AtCBL4 interact specifically with CIPK-A

To test for interactions between the CBLs and CIPKs the genes for each CBL and CIPK had to be cloned from *M. polymorpha*. As an additional test AtCBL4 (SOS3) and AtCIPK24 (SOS2) were also cloned to assess for interactions between the CBLs and CIPKs from *M. polymorpha*. AtCBL4 and AtCIPK24 interaction also acts as a positive control for the interaction test as it has already been demonstrated in yeast two-hybrid (Halfter et al. 2000). As a negative control both CIPK-A and CIPK-B were cloned with the 24 amino acid NAF domain removed and labelled CIPK-AΔNAF and CIPK-BΔNAF as the NAF domain is required for CBL-CIPK interactions. All of these CBL and CIPK genes were cloned into a split GAL4 system for yeast two-hybrid tests with the CBLs fused to the GAL4 binding domain (CBL-BD) and the CIPKs fused to the activation domain (CIPK-AD).

Based on the yeast two-hybrid CBL-B and CBL-C can interact with AtCIPK24 and therefore they may be potential SOS3 (AtCBL4) homologues (Fig.5.1). CBL-B has a strong interaction in the yeast two-hybrid with CIPK-A and AtCIPK24 with survival of colonies on the selection plate to five serial dilutions (1:10). CBL-B also interacts with CIPK-B although serial dilution suggest that it may interact less strongly (Fig.5.1). CBL-C interacts strongly with both CIPK-A and AtCIPK24 but does not seem to interact with CIPK-B in yeast two-hybrid (Fig.5.1). CBL-A interacts strongly with CIPK-A and CIPK-B but not with AtCIPK24 and therefore is not likely a potential SOS3 (AtCBL4) homologue (Fig.5.1). Based on the yeast two-hybrid either CBL-B or CBL-C could be a SOS3 (AtCBL4) homologue or they could be functionally redundant and both be SOS3 (AtCBL4) homologues.

A reciprocal test was used to check for MpCIPK interactions with AtCBL4 to determine the potential SOS2 (AtCIPK24) homologue. AtCBL4 interacted strongly with AtCIPK24 as a positive control and also interacted with CIPK-A which sug-

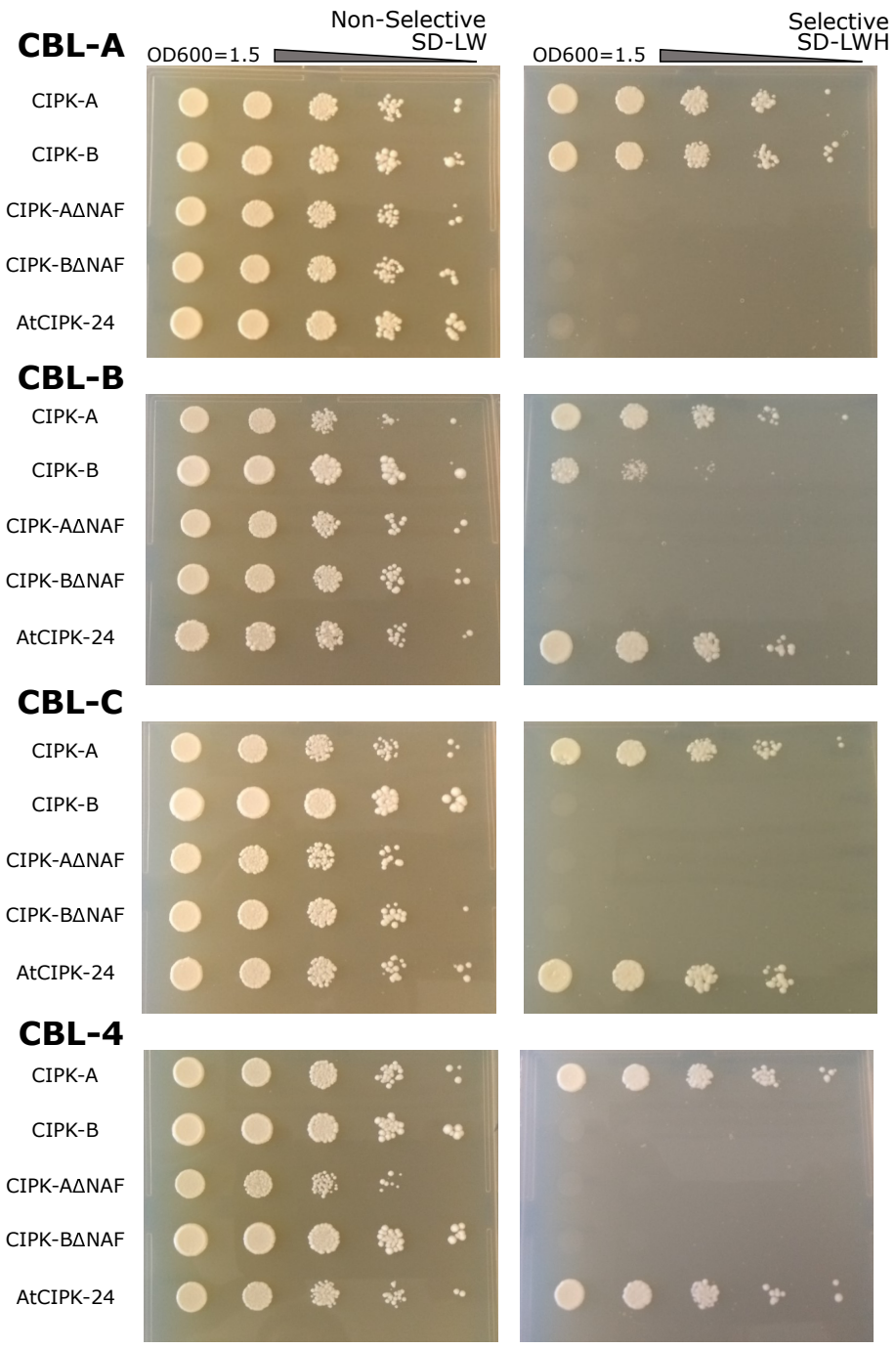


Figure 5.1: Yeast Two-Hybrid to test for interactions between CBLs and CIPKs from *M. polymorpha*. CBLs fused to the GAL4 BD transformed into yeast (Y187) and CIPK's fused to the GAL4 AD transformed into yeast (AH109) which are mated into different combinations. Yeast are grown to OD600=1.5 and serial diluted (1:10) and plated on growth plates (SD-LW) and interaction test plates (SD-LWH). Three CBLs from *M. polymorpha* (A/B/C) and one *A. thaliana* (4) were tested against 2 CIPK's from *M. polymorpha* (A/B) and one from *A. thaliana* (24) with negative controls including CIPK's from *M. polymorpha* with NAF interaction domain removed.

gests it may be a SOS2 homologue (Fig.5.1). AtCBL4 did not interact with CIPK-B suggesting that it is not a SOS2 homologue. None of the CBLs interacted with

CIPK-AΔNAF or CIPK-BΔNAF and therefore the interactions with CIPK-A and CIPK-B determined in the yeast two-hybrid are dependent on the NAF domain in the MpCIPKs as would be expected, therefore false positive interactions are unlikely (Fig.5.1). Hence, MpCIPK-A is likely to be the SOS2 homologue in *M. polymorpha* based on yeast two-hybrid results.

Western blot was used to confirm the expression of all of the proteins in the yeast to ensure that none of the negative interactions were caused by lack of expression. CBLs were fused with the binding domain of the GAL4 transcription factor with a c-myc tag and therefore blotting was carried out with a c-myc antibody and protein was detected with the yeast native peroxidase PRX1 used as a loading control (Fig.5.2). All MpCBLs and AtCBLs are detectable by western blot and therefore all of the CBLs were expressed in the yeast two-hybrid. CIPKs were fused with the activation domain of the GAL4 transcriprion factor with a HA tag and therefore blotting was carried out with anti-HA and protein product was assesed (Fig.5.2). All the CIPKs could be detected by western blot through the HA tag and therefore all of the CIPKs were expressed in yeast two-hybrid. As all of the protein products for the CBLs and CIPKs could be detected by western blot it is unlikely that any false negative interactions happened because of lack of expression.

5.2.2 Bimolecular Fluorescence Complementation determined that all MpCBLs and MpCIPKs can interact in *N. benthamiana*

Yeast is very different from *M. polymorpha* and does not contain the same subcellular localisation sites including the vacuole. CBL-A was predicted to localise to the vacuole and lack of this localisation site may affect the ability for it interact with CIPKs. Yeast also has different mechanisms for some post translational modifications. CBL-B/C are thought to be posttranslationally modified through myristoylation and palmitoylation and, while the mechanism is conserved in yeast, the localisation relies on the MGCxxS motif being N-terminus to the rest of the pro-

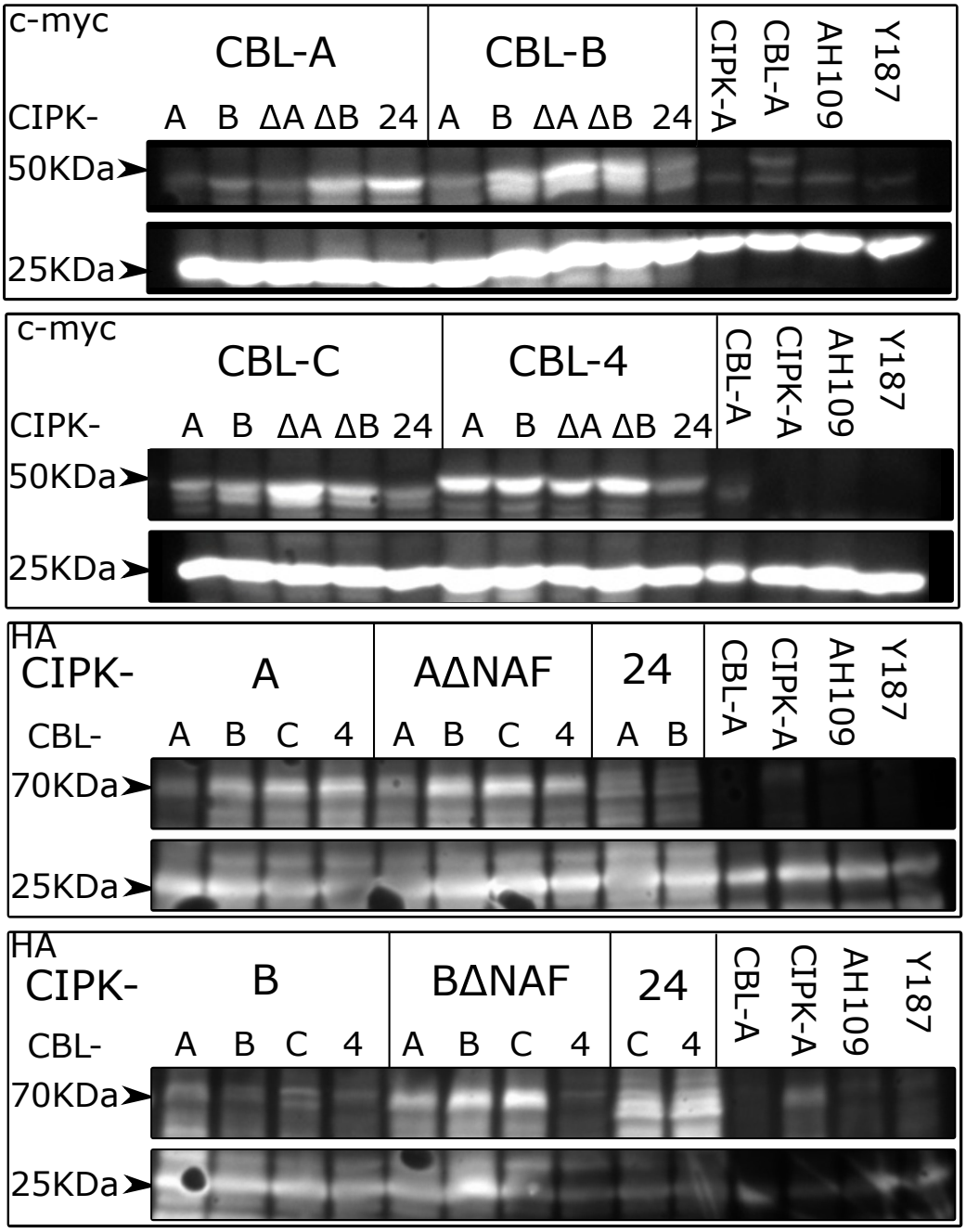


Figure 5.2: Western Blot for CBLs and CIPKs from Yeast Two-Hybrid. Western blot for c-myc tagged CBLs and HA tagged CIPK's for strains used in the Yeast Two-hybrid. Expected sizes for the CBLs fused to GAL4 binding domain and c-myc tag with expected protein sizes of 48.49 kDa (MpCBL-A), 46.63 kDa (MpCBL-B), 46.56 kDa (MpCBL-C) and 47.75 kDa (AtCBL-4). Expected sizes for CIPKs fused to GAL4 activation domain and Ha tag for predicted protein sizes of 69.71 kDa (MpCIPK-A), 69.94 kDa (MpCIPK-B), 67.11 kDa (MpCIPK-AΔNAF), 67.37 kDa (MpCIPK-BΔNAF) and 70.17 kDa (AtCIPK-24). Native peroxidase activity (PRX1) was used as a loading control with and expected size of 25 kDa.

tein which it would not be as the GAL4-BD was fused to the N-terminus. Therefore it was important to confirm interactions in a less heterologous system using a plant

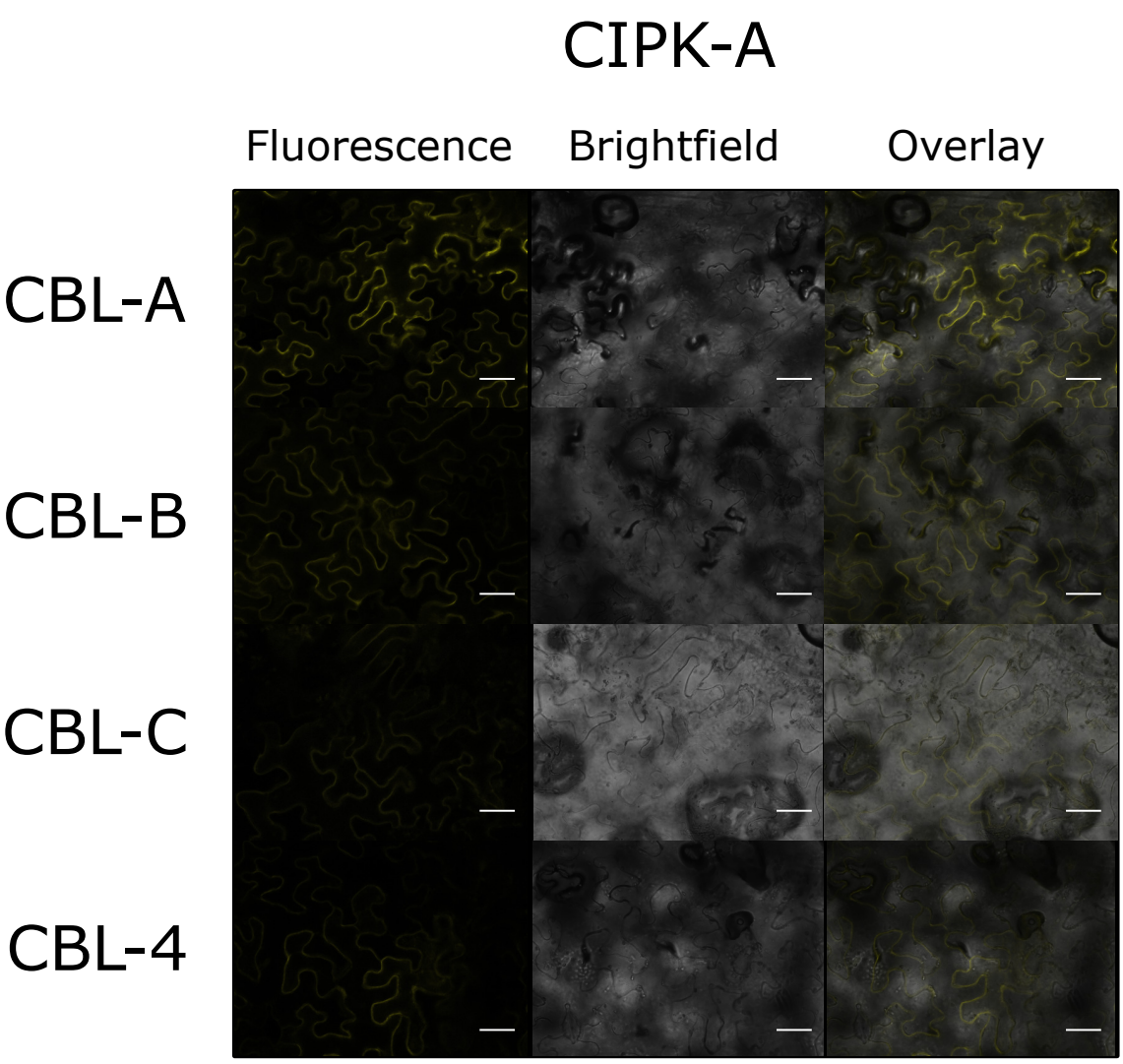


Figure 5.3: Bimolecular fluorescence complementation to assess interactions between MpCBLs (A/B/C), AtCBL-4 and CIPK-A. BiFC includes fluorescence, brightfield and overlay images to assess for interaction and localisation of the interaction through transient transformation of *N. benthamiana*. White bars represent 50 μm.

system which has the same subcellular localisation sites as *M. polymorpha*. Bimolecular fluorescence complementation (BiFC) utilises a split YFP with each part of the YFP fused to a seperate protein which are tested for interaction and can be used in plant systems such as *Nicotiana benthamiana*. The YFP protein parts have low affinity and therefore if interaction takes place it will result in fluorescence signal that is not detectable if the proteins of interest do not interact. Previous studies have fused the VYNE to CBLs on the C-terminus and found interaction with VYCE fused to the N-terminus of the CIPK which would avoid mislocalisation. Therefore, bimolecular fluorescence complementation constructs were designed with

CBLs fused to VYNE (CBL-VYNE) and CIPKs fused to VCNE (VYCE-CIPK) and transiently transformed in *N. benthamiana*.

Bimolecular fluorescence complementation (BiFC) confirmed interactions between CBL-A/B/C from *M. polymorpha* and AtCBL4 with CIPK-A. Therefore, Transient transformation of *N. benthamiana* with BiFC constructs demonstrated potential interaction with all the MpCBLs and AtCBL used in the study which confirms the result found in the yeast two-hybrid. Fluorescence does not determine interaction strength due to lack of expression control for the construct. However all the CBLs seem to interact in 4 technical replicates used in this study with MpCBL-C and MpCIPK-AΔNAF tested as a negative control with no fluorescence under the same conditions (data not shown). The potential interaction with all the CBLs does show a lack of specificity in interactions with CIPK-A in *M. polymorpha* and therefore it may be possible that there is some functional redundancy in the system or that three potential pathways exist for CIPK-A with each CBL. The interaction with AtCBL4 also may imply that CIPK-A is a AtCIPK24 (SOS2) homologue in *M. polymorpha* although this does not necessarily mean it operates in the SOS-dependent salt pathway.

Interactions were also confirmed with CIPK-B from *M. polymorpha* using BiFC constructs transiently transformed into *N. benthamiana* and confirmed interaction with CBL-A and CBL-B, but also showed an interaction with CBL-C and AtCBL4 with CIPK-B. CBL-C and AtCBL4 interaction with CIPK-B was not detectable in the yeast two-hybrid test and therefore the result of the BiFC contrast the result shown in yeast. As CBL-C and AtCBL4 are myristoylated and palmitoylated alongside CBL-B it could be that the lack of this posttranslational modification interferred with interaction in yeast but this is unlikely as CBL-B could still interact unless the nature of the interaction differs between these CBLs. Alternatively the GAL4 BD may have interferred in this interaction but this should also effect interaction with the other CBLs unless CBL-A/B have a different interaction to CBL-C and AtCBL4. Once again 4 technical replicates were used to assess the interaction and therefore it

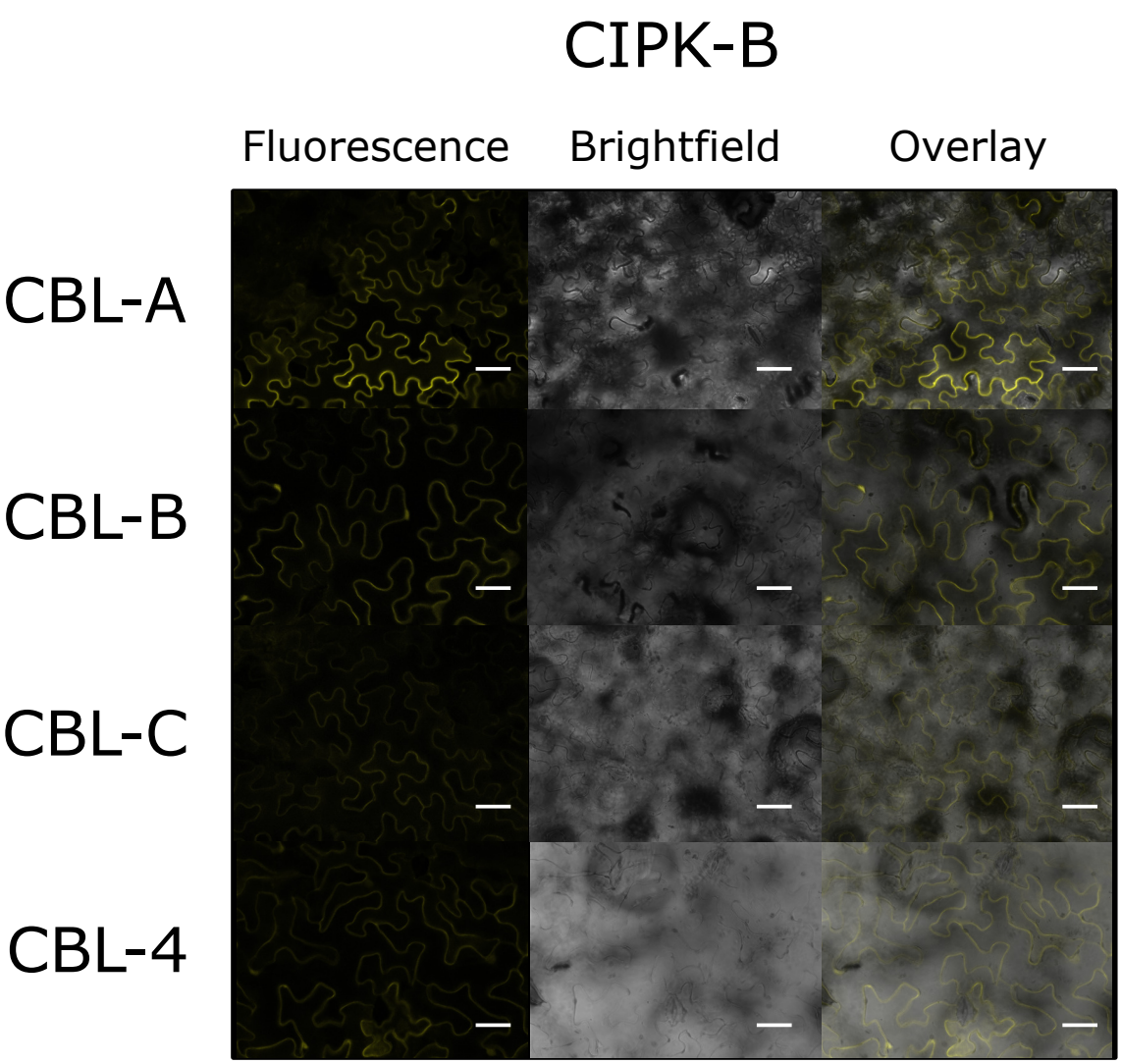


Figure 5.4: Bimolecular fluorescence complementation to assess interactions between MpCBLs (A/B/C), AtCBL-4 and CIPK-B. BiFC includes fluorescence, brightfield and overlay images to assess for interaction and localisation of the interaction through transient transformation of *N. benthamiana*. White bars represent 50 μm.

is a consistent result. As the AtCBL4 demonstrated no specificity between CIPK-A/B it could not be conclusively determined that either are a SOS2 homologue. The potential for interactions between all of the CBLs and CIPKs from *M. polymorpha* also demonstrated a lack of specificity between CBL-CIPK interactions which may imply either diverse signalling responses through different interaction or may give more evidence for the possibility of functional redundancy.

As the CIPKs from *M. polymorpha* did not demonstrate any specificity in interactions with the CBL-A/B/C or AtCBL4, interactions were also tested between the CBLs and AtCIPK24 as AtCIPK24 is known to be selective in interactions in *A.*

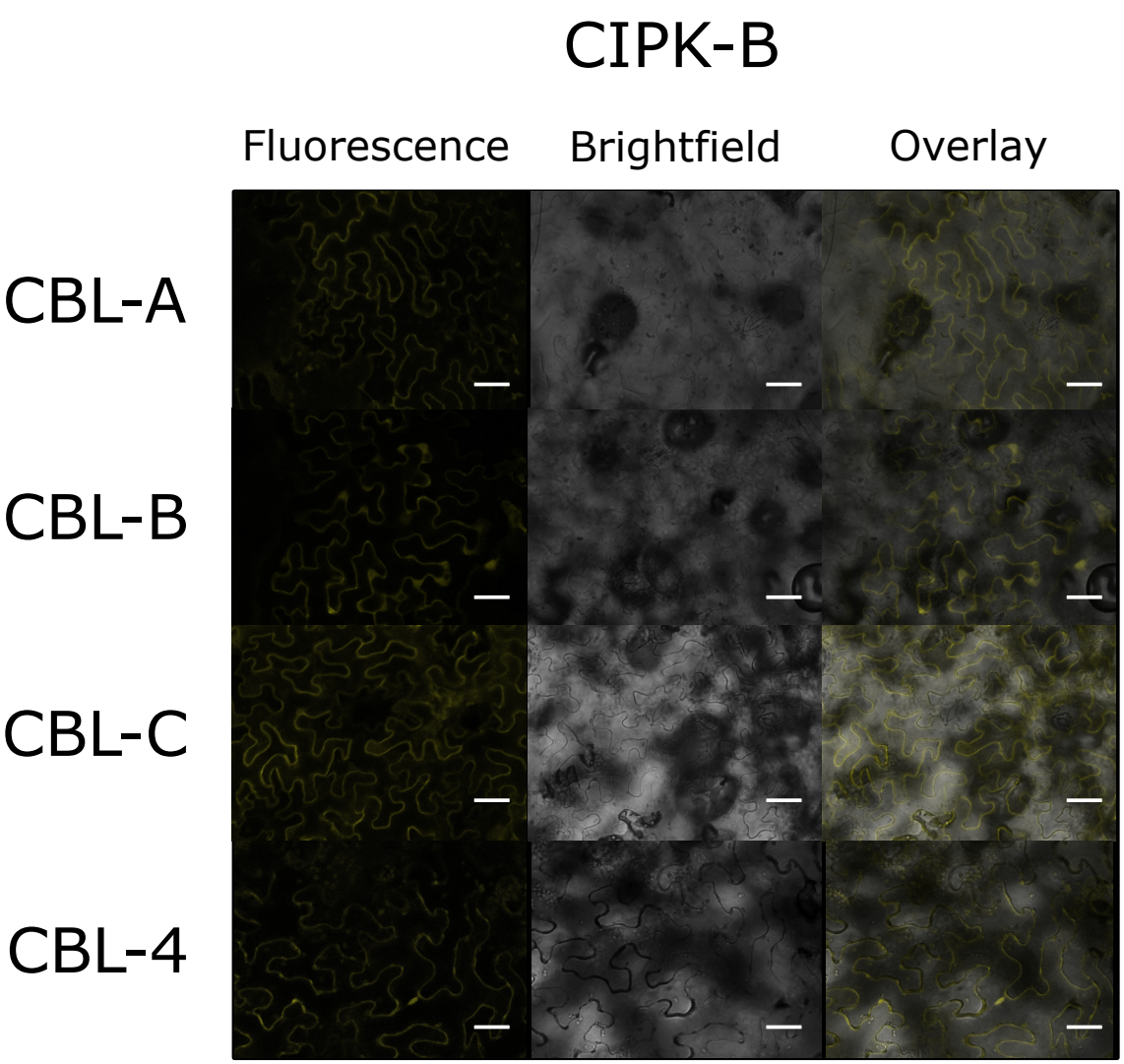


Figure 5.5: Bimolecular fluorescence complementation to assess interactions between MpCBLs (A/B/C), AtCBL-4 and CIPK-24. BiFC includes fluorescence, brightfield and overlay images to assess for interaction and localisation of the interaction through transient transformation of *N. benthamiana*. White bars represent 50 μm.

thaliana. BiFC confirmed interaction between AtCIPK-24 and CBL-A/B/C as well as AtCBL4. AtCBL4 was already a known interaction as it is SOS3 and known to interact with SOS2 (AtCIPK24) in yeast two-hybrid, BiFC and in *in vitro* Co-IP experiments (Halfter et al. 2000; Quan et al. 2007). The yeast two-hybrid experiment demonstrated interaction between CBL-B/C and AtCIPK24 and these interactions were conserved in the BiFC. However CBL-A was not demonstrated to interact with AtCIPK24 in yeast and as it is the only CBL from *M. polymorpha* thought to localise to the tonoplast it may be that this potential for localisation is required for it to be able to interact with AtCIPK24. Based on the images of the interaction

between CBL-A and AtCIPK24 it is difficult to say whether CBL-A is localised to the tonoplast although it does seem less localised to the plasma membrane than the interaction between AtCBL4 and AtCIPK24. A fluorescence control known to localise to the tonoplast would be needed to demonstrate the same localisation in CBL-A. Alternatively it could be that all CBLs and CIPKs are capable of interacting in *M. polymorpha* and perhaps this is true for all ancient CBLs and it was only through evolution of diverging structures that led to only certain subset being capable of interacting in *A.thaliana*

5.3 Discussion

Based on the Yeast two-hybrid result it could be proposed that MpCBL-B or MpCBL-C are SOS3 homologues in *M. polymorpha* due to the interaction with AtCIPK24. However, there are two potential CBLs from *M. polymorpha* which mean they are functionally redundant, or they could also fulfill other roles. It is perhaps more likely that they fufill multiple roles as CBL-B also interacts with multiple MpCIPKs. However, AtCIPK24 has been demonstrated to show interactions with other AtCBLs including AtCBL10 (ScaBP8/SOS4) which is expressed in the shoot of *A. thaliana* to also take place in salt responses. Similarly, AtCIPK24 interacts with AtCBL1 in yeast two-hybrid which is induced in expression in different time frames in response to drought, cold and high salinity and knockouts of *AtCBL1* results in sensitivity to drought, cold and high salinity (Kolukisaoglu et al. 2004; Albrecht et al. 2003; Cheong et al. 2003). Alongside this AtCBL9 also interacts with AtCIPK24 in yeast two-hybrid and *AtCBL9* knockout is known to be involved in a number of ABA responses including drought responses although none of these responses are currently ascribed to interaction with AtCIPK24 (Kolukisaoglu et al. 2004). AtCIPK24 is also found to be involved in phosphorylations and increased activation of a Ca^{2+}/H^{+} antiporter CAX1 with no currently ascribed CBL partner (Cheng et al. 2004). As AtCAX1 is associated with cold tolerance it could be that AtCBL9 is involved in this pathway but that is unlikely as AtCBL9 is currently described

as localised the plasma membrane and CAX1 activity takes place at the tonoplast unless AtCBL9 relocalises or AtCIPK24 can seperate and still retain active confirmation which is not currently known (Catalá et al. 2003). Therefore MpCBL-B/C could be SOS3 homologues as it would be expected that they can still interact with a number of other CIPKs. As AtCIPK24 is known to interact with a number of CBLs it may also be that they represent two different pathways with CBL-B acting as a homologues of AtCBL1/9. However in the case of AtCIPK-24 it is particularly interesting that even when interacting with other CBLs many of the responses are still involved in high salinity or drought responses.

AtCBL4 was also included in the yeast two-hybrid to try and determine if there was a SOS pathway in *M. polymorpha*. The interactions between CBLs and AtCIPK24 implied the CBL-B/C as potential SOS3 homologues and the interactions between AtCBL4 and the CIPKs may also imply which CIPK could be a SOS2 (AtCIPK24) homologue. Based on the yeast two-hybrid AtCBL4 was found to specifically interact with CIPK-A of the *M. polymorpha* CIPKs. CBL-C also specifically interacted with CIPK-A and AtCIPK24 which implied a high likelihood of MpCBL-C as a AtCBL4 homologue. AtCBL4 has been shown to not interact with AtCIPK1-16 (Zhang and Lu 2003). However a study also demonstrated an interaction between AtCBL4 and AtCIPK1 using the same vectors and yeast strain (Shi et al. 1999). Another group also demonstrated interaction between AtCBL4 and AtCIPK6 for plasma membrane targeting of K^+/H^+ antiporter AKT2 for K^+ homeostasis (Held et al. 2011). All of the proteins were confirmed to be expressed through western blot and removal of the NAF domains from MpCIPKs did not show an interaction which would imply that there were no false negatives however the selectivity of the interactions could not be confirmed with bimolecular fluorescence complementation.

Bimolecular fluorescence complementation carried out in *N. benthamiana* could determine no specificity in interactions between CBLs and CIPKs from *M. polymorpha* with MpCBL-C and MpCIPK-AΔNAF tested with no fluorescence under the

same conditions (data not shown). As the test relies on both of the proteins containing half of a fluorescence tag are in close proximity to generate fluorescence, the presence of fluorescence demonstrates the proteins were expressed. As all constructs demonstrated fluorescence it could be that overexpression of the constructs caused proteins that would typically not interact, to interact. However a negative control was tested by generating a CIPK-AΔNAF BiFC construct and checking interaction with CBL-A and no fluorescence could be detected in two technical replicates. Technically it could be that the CIPK-AΔNAF did not express so it may be optimal to repeat this experiment and also check for the protein product (e.g. by western blot). Another test that would supplement this result would be to check tissue specific expression of the CBLs and CIPKs using native promoters with the CIPKs fused to GFP, as it may be that some CBL-CIPK interaction can take place but do not *in vivo* as the proteins are not expressed in the same tissue. GFP fused CIPKs could also be used for co-immunoprecipitation to assess both tissue specificity and which proteins are interacting in the same experiment. However based on the results thus far we could not determine any specificity in *M. polymorpha* CBL-CIPK interactions through BiFC in *N. benthamiana*. One possibility is that the different combinations of protein interaction (e.g. CBL-B/CIPK-A and CBL-C/CIPK-A) do take place *in vivo* but activate different downstream responses under different calcium signals. This would be similar to the case where AtCIPK24 can form an interaction with AtCBL4 for the SOS pathway or AtCIPK-24 can interact with CBL1 for a seperate pathway. If all interactions are possible and all interactions activate different pathways then there could be up to 6 different pathways of response for CBLs and CIPKs in *M. polymorpha*.In comparison to this *P. patens* has more selectivity with 19 of 28 interactions possible by yeast two-hybrid however more may be possible with BiFC (Kleist et al. 2014). Alternatively, some of the CBL-CIPK interactions could be functionally redundant and more than one CBL can interact with a single CIPK to activate the same pathway. This cannot be determined from interaction tests but should be considered.

Differences between the constructs used to generate the yeast two-hybrid plasmids could have caused some of the interaction differences between yeast and tobacco interaction experiments. Following previous experiments to test yeast two-hybrid interactions between CBLs and CIPKs, the CIPKs were fused to the activation domain of GAL4 on the N-terminus of the protein, and the CBLs were fused with the binding domain of GAL4 on the N-terminus of the protein (Halfter et al. 2000; Zhang and Lu 2003; Hashimoto et al. 2012). In contrast the BiFC experiments were carried out with the with VYNE fused to the C-terminus of the CBL and VYCE fused to the N-terminus of the CIPK as per previous publications (Hashimoto et al. 2012; Waadt et al. 2008). The technical reason for adding the fusion to the C-terminus of the CBL is that localisation is determined by the N-terminus of the protein with myrisoylation motif (MXXXS) containing a palmitoylation site hence (MGCXXS) for dual fatty acid localisation to the plasma membrane, which allows for localisation to be determined in BiFC alongside interaction. However this was not considered important in previous yeast two-hybrid experiments but yeast can also myristoylate and palmitoylate proteins with an N-terminus motif (MGCXXS; Johnson et al. (1994); Greaves et al. (2009)). While the localisation itself is not thought to affect interaction the dual acid modification could affect the tertiary structure of the protein and interfere with potential protein:protein interactions.

The NAF domain of a CIPK is necessary and sufficient for interaction with with a CBL but the binding pocket of the CBL is less well defined and seems to vary more between CBLs. It may be that the myristoylation could be required for proper conformation to allow from protein:protein interaction or it may be required for interaction as part of the interface. Previously it has been demonstrated that myristoylation is required for interaction between CAP-23/NAP-22 and calmodulin and myristoylation is required for phosphorylation of PKC in rat brains and in *in vitro* interaction tests with proteins purified from *E. coli* (Sánchez-Barrena et al. 2005). SOS3 has also been demonstrated to require N-myristoylation to confer salt tolerance but it was thought that the myristoylation was required for targeting to

downstream SOS1 rather than because it affects interaction with SOS2. However, no change in localisation could be detected in SOS3(G2A) mutation and therefore it may be the lack of myristoylation actually affects protein interaction with SOS2 (Ishitani et al. 2000). Although, AtCIPK24 and AtCBL4 still interact in the yeast two-hybrid and therefore it may be that AtCBL4 is not modified in myristoylation attachment or it may be that the myristoylation is not required for interaction. Other explainations could include that yeast two-hybrid results in different calcium occupancy which could affect protein interactions between CBLs and CIPKs as calcium occupancy is well known to affect protein interactions between certain CIPKs and CBLs (Batistič and Kudla 2004; Luan et al. 2002). Similarly it may simply be that the GAL4 tag or VYNE tag interferes with proper conformation of the CBL which may have an impact on protein interactions which is why multiple interaction tests should be used.

Chapter 6

Knockout of *CIPK-B* in *M. polymorpha*

6.1 Genetics of signalling in *M. polymorpha*

A number of basic signalling processes are conserved in *M. polymorpha* but with reduced numbers of protein components. A few key differences between *M. polymorpha* and angiosperms have already been discovered and understanding these differences may provide insight into how these signalling pathways evolved to have more functional redundancy and complexity. For example, MpPIF has been described as a part of light based phytochrome signalling in *M. polymorpha* but with only a single PIF protein (Inoue et al. 2019). MpPIF has also been demonstrated to be involved in a conserved MpDELLA signalling for growth control in response to modulate developmental traits (Hernández-García et al. 2021). However a key difference is that, while MpDELLA is present, gibberellic acid (GA) signalling components have not been found and MpDELLA activity seems to act independent of this process in *M. polymorpha*. However, MpDELLA can still respond to oxidative stress and other stress conditions which implies that the DELLA activity predates GA signalling in plants (Hernández-García et al. 2021). Similarly jasmonate signalling components have been discovered and it has been demonstrated that the jasmonate-isoleucine

precursor ODPA is responsible for activity in *M. polymorpha* with MpCOI1 as a receptor and a single MpJAZ repressor (Monte et al. 2019). While *M. polymorpha* can also respond to ethylene, it cannot synthesise ethylene, and instead produces the precursor 1-aminocyclopropane-1-carboxylic acid (ACC) and responds in a similar manner to exposure to either of these plant hormones (Monte et al. 2019; Li et al. 2020). However, both auxin signalling and cytokinin signalling is conserved but with fewer genes involved in *M. polymorpha* (Flores-Sandoval et al. 2015; Kato et al. 2015; Aki et al. 2019). Therefore a lot of the key components are present for plant signalling, and many of the plant hormonal signalling components have been described. These studies have demonstrated that the signalling components are similar to angiosperms and that they have a reduced number of proteins constituting the signalling pathways. Some of this signalling has been demonstrated to be different, and uses precursors for plant hormones, instead of established hormones, such as OPDA instead of jasmonate-isoleucine, and ACC instead of ethylene. These finding can further studies to elucidate the evolution of these ancient hormonal signalling components. However, it first has to be established if these signalling pathways do represent ancient signalling functions or if there is a loss of jasmonate-isoleucine and ethylene production in *M. polymorpha*.

6.2 Molecular tools for genetic manipulation of *M. polymorpha*

During and prior to these studies a range of molecular biology and genetic tools have been developed, such as tools for transformation of different tissues of *M. polymorpha*. Initial transformation of *M. polymorpha* was carried out through particle bombardment (Irifune et al. 1996). Since then a range of tools have allowed for *Agrobacterium*-mediated transformation of a range of tissues in *M. polymorpha*. Some of the first transformations were carried out on immature thalli grown from sporelings which did not required high transformation efficiency due to the number

of spores that can be transformed. The *Agrobacterium*-mediated transformation of sporelings require 100 μM acetosyringone, 7 days of growth for immature thalli and co-culture with *Agrobacterium* for 3 days (Ishizaki et al. 2008). Since then higher transformation efficiencies have been established using techniques for transformation of mature thallus tissue by incubating the apical-basal axis of cut thalli (Kubota et al. 2013). Subsequently, the difficulty of transformation has been reduced further using single plate methods of transformation called S-AgarTrap for sporeling, T-AgarTrap for thallus, and G-AgarTrap for gemmae transformation (Tsuboyama-Tanaka and Kodama 2015; Tsuboyama and Kodama 2014; Tsuboyama et al. 2018). The use of these techniques reduces resources required for transformation such as inducing sexual structure formation and eases the difficulty of replicating techniques for stable transformation of *M. polymorpha*. Very recently transient transformation has been demonstrated for *M. polymorpha* which opens a wide range of possibilities for testing localisation and interactions between proteins in *M. polymorpha* (Iwakawa et al. 2021).

The development of transformation technologies allows for further molecular biology techniques such as genetic knockout to investigate gene function. CRISPR/-Cas9 is a gene editing technology which requires a single-guide RNA (sgRNA) to target a section of the genome and a Cas9 enzyme which is directed by the guide RNA to cause a double strand break in DNA which is then repaired inaccurately by non-homologous end joining (NHEJ), or can be directed to give precise mutation through homology-directed repair (HDR). CRISPR/Cas9 has been successfully used in *Oryza sativa, Triticum aestivum, Nicotiana benthamiana, Sorghum bicolor* and *Arabidopsis thaliana* to investigate gene function in flowering plants (Xie and Yang 2013; Li et al. 2013; Nekrasov et al. 2013; Jiang et al. 2013; Feng et al. 2013; Mao et al. 2013). More recently the technology has been deployed in bryophytes such as *M. polymorpha* and *Physcomitrella patens* (*Physcometrium patens*) (Sugano et al. 2014; Lopez-Obando et al. 2018). Alongside this, gemmae propogation can be used to continue isogenetic lines and tools have also been developed for the storage of

these lines which allows for sharing of genetic lines between laboratories. The first cryopreservation protocol required dehydration on high sucrose media followed by encapsulation in alginate beads for storage (Tanaka et al. 2015). A simplified version of the protocol has since demonstrated that encapsulation is not required but just the addition of sterile silica beads to the tube of gemmae. However this protocol is more effective on gemmae of 10-33 days old which is not a consideration when gemmae are encapsulated (Takahashi and Kodama 2020). The development of both cryopreservation and gene knockout technology in *M. polymorpha* is key to be able to carry out gene function analysis in this system.

CRISPR/Cas9 gene editing in plants has gone through multiple optimisations for use in plants. The Cas9 enzyme has been codon optimised and tested in tobacco hence called plant codon optimised Cas9 (pco-Cas9). pco-Cas9 has had introns introduced from potato to stop erronous expression and activity during cloning in bacteria (Li et al. 2013). Similarly, it has been found that for the best editing outcomes the sgRNA requires endogenous U6 promoters for DNA polymerase III expression (Long et al. 2018; Sun et al. 2015). An endogenous U6 promoter has already been described in the inital testing of CRISPR/Cas9 editing in *M. polymorpha* (Sugano et al. 2018). Under this system U6 has been demonstrated to have higher expression if the first base of the sgRNA has a G or A as the first base and for Cas9 recognition the sgRNA requires a PAM motif and must end in an 'NGG' (Hassan et al. 2021). For efficient gene editing the expression of Cas9 also has to be considered and the p35S constitutive promoter has been demonstrated to not have activity in all parts of the plant (Kanazawa et al. 2016) hence the native MpEF1α is better to drive expression as it is expressed in all parts of the plant (Althoff et al. 2014). Importantly it was also demonstrated that efficiency of CRISPR/Cas9 activity is improved with the addition of a plant nuclear localisation sequence (Sugano et al., 2018). With the development of effective promoters, optimised genes and endogenous U6 for guide preparation CRISPR/Cas9 gene editing is efficient to knock out genes of interest if the sgRNA can be effectively targetted to the gene of interest.

6.3 Investigating function of *CIPK-B* through genetic knockout

To decode the functions of the CBL-CIPK network in *M. polymorpha* it is possible to investigate function through specific genetic knockout of the genes involved. Specific knockout of MpCIPK-A is hypothesised make *M. polymorpha* plants that are more sensitive to high concentrations of salt due to the interactions with MpCBL-C in *Saccharomyces cerevisiae*. *MpCBL-C* is modified in expression in response to salt treatment in a dosage dependent manner and therefore is the most likely CBL to be involved in salt tolerance in *M. polymorpha*. MpCIPK-B does not have a proposed function based on the analysis presented before (Chapter 5.1.3) but it could also be the salt sensitive MpCIPK as interaction tests in *N. benthamiana* demonstrated that either MpCIPK can interact with MpCBL-C (Fig.5.4). However, it could also be that either of the MpCIPKs could be involved in other functions such as dealing with heavy metal contamination, potassium uptake, or drought conditions. Functions of CBLs and CIPKs in *A. thaliana* also overlap with salt stress responses such as drought due to the osmotic effect of salt. K^+ uptake is also involved in salt tolerance as K^+/Na^+ balancing is common in plants. Na^+ chelates phosphates and is therefore toxic to cells and also can depolarise membranes in high concentrations which blocks K^+ uptake and results in very low K^+ concentration in plants. This creates a malnutrition in the plant for K^+ which could trigger uptake of K^+ from the environment to rebalance these ions. Therefore, heavy metal sensitivity may be the best phenotype to assess for functions not related to salt when initially trying to characterise CIPK knockout mutants.

AtCIPK23 responds to heavy metal accumulation in *A. thaliana* and therefore there may be a homologue of this present in *M. polymorpha*. AtCIPK23 interacts with and phosphorylates IRT1 which is involved in iron uptake and non-selective uptake of other heavy metals such as zinc, manganese, cobalt and cadmium (Dubeaux et al. 2018). With increased concentration of metals other than iron, IRT1 is inter-

nalised to the vacuole instead of its usual plasma membrane localisation. Alongside this internalisation, IRT1 is degraded by an ubiquitination pathway initiated by the phosphorylation of IRT1 by AtCIPK23 (Dubeaux et al. 2018). Without AtCIPK23 plants cannot remove IRT1 and therefore continue to accumulate heavy metals and hence shown an enhanced susceptibility to these metals and a toxicity response including chlorosis and perturbed growth (Ares et al. 2018; Dazy et al. 2009). Typically plants are very sensitive to high concentrations of heavy metals but previous studies have demonstrated that *M. polymorpha* can survive 2 mM $ZnCl_2$ for three weeks (Ares et al. 2018) and therefore higher concentrations can likely be tolerated for 7 days in the wildtype. Mutants that interfere with IRT1 degradation are likely to be more susceptible to high zinc and therefore we can expect increased susceptibility either earlier or at lower concentrations. Preliminary studies in the Miller group demonstrated that wildtype Cam2 does not survive 14 days with 5 mM $ZnSO_4$ and therefore this can be used as a higher concentration for assessment at 7 days (James Houghton, Personal Communication).

If the functions of MpCIPK-A/B can be defined, then the range of functions that the CBL-CIPK pathways are involved with in *M. polymorpha* can be supposed. Knockout of MpCIPKs should show increased susceptibility or tolerance to the stress that the proteins are involved in, provided there is no functional redundancy between the proteins. Knockout of MpCBLs can then be generated and compared to the MpCIPK functions to assess for overlap which, including the interaction data (Chapter 5.2), can be used to deconvolute many of the functions of the pathways. The functions of the MpCIPKs may represent the original functions of CBL-CIPKs in early land plants. Simultaneously novel functions of the CBL-CIPK pathway may also be found in *M. polymorpha* through broad screening of different phenotypes.

6.4 Results

6.4.1 CRISPR/Cas9 Constructs and Transformation of *M. polymorpha*

CRISPR/Cas9 constructs were designed with a hygromycin selection casette (p35S-*hpt*-t35S) for selection of positive transformation by placing plants on hygromycin. The Cas9 enzyme pcoCas9 was added to the construct to induce double strand breaks in the genome based on guidance of sgRNA with included potato introns to stop expression in *E. coli* and hence potential erronous breaks in the construct. The Cas9 enzyme also had an additional nuclear localisation sequence (Lit6b) to direct the enzyme to the nucleus for efficient double strand breaks. Two single guides were designed for the first exon of the genes of interest. The use of two sgRNAs has the benefit of removing of a section of the gene which would allow for screening of mutants with PCR. Additionally the design of two single guides helps to account for the variable efficiency of different single guides in the genome. If hygromycin resistance is detected in transformed plants and the PCR product does not have the section between the sgRNA target sites removed it may be that one of the singleguides has a low efficiency, and therefore a section cannot be removed, but the other is still effective and therefore the double strand break could only happen in a single location. This case could be detected through sequencing of the PCR product to detect if a mutation is caused, and if that mutation causes a frameshift which would still make a gene knockout. Single guides were expressed with a RNA polymerase III through the *M. polymorpha* U6 promoter. The full construct design can be found in the methods (Chapter 2.1.2).

T-Agar Trap and G-Agar Trap protocols could not be replicated in our labratory after 12 replicates per protocol. Each of the attempts used 100 plantlets per replicate with thallus cuttings as directed for T-Agar Trap and gemmae for G-Agar Trap. Therefore at least 1200 plants were used per protocol and no positive transformants could be detected for either. Fertilised sporeheads were donated

by the Haseloff Labratory (Cambridge) and the Sporeling Transformation protocol was checked which allowed for assessment of the CRISPR constructs. MpCDPK-C was knocked-out but the mutants were found to have a development mutation, and could not be effectively grown to determine the nature of the growth defect, but did demonstrate that the CRISPR constructs were correctly designed (data not shown). Preliminary studies in the Miller group used a GUS expression construct under a constitutive expression promoter (p35S) to try to modify the sporeling protocol to be effective on the gemmae and found that using MES-$MgCl_2$ infiltration buffer with *Agrobacterium tumefaciens* GV3101 of an OD_{600}= 0.5 could be used to transform plants with an efficiency of 0.3% (Elizabeth Payne, Personal communication). Therefore the modified sporeling transformation protocol was to generate CRISPR/Cas9 mutants in this study (Chapter 2.1.5) through transformation of gemmae.

6.4.2 Guide design and Genotyped lines

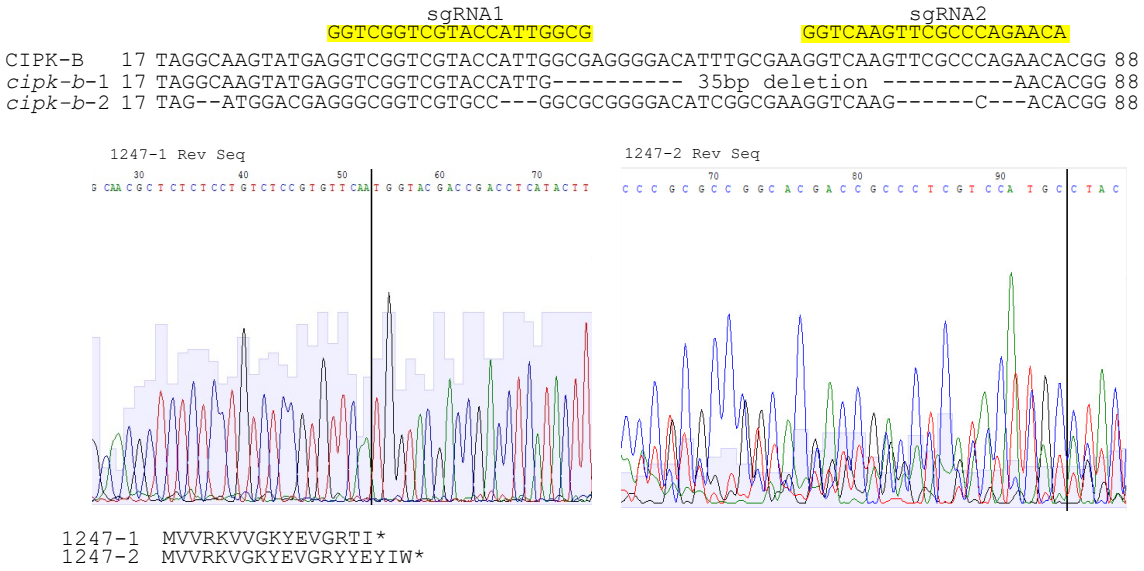


Figure 6.1: sgRNA design for CRISPR/Cas9 of CIPK-B including sequenced *cipk-b* lines and predicted protein sequences of mutants. Exon 1 of CIPK-B from *M. polymorpha* with the target of the two RNA single guides labelled. Generated lines of CIPK-B mutants gene sequence is shown in comparison to wild type along with images of the trace files of the sequencing through Sanger sequencing. Predicted protein sequence of mutants were predicted with Expasy Translate to show the gene deletion product.

The first exon of each of the MpCIPKs were sequenced from Cam2 gDNA to check if it matches the sequence of the Tak1 genome (v5.1, marchantia.info). MpCIPK-A of Cam2 had a single nucleotide polymorphism compared to the genome which was located in the second single guide which was designed. MpCIPK-B of Cam2 matched the sequence in the genome and therefore no polymorphisms were detected. Both constructs for MpCIPK-A and MpCIPK-B knockout were transformed and two lines of MpCIPK-A mutants were found but were not be assessed within the scope of this project. Several lines for MpCIPK-B were detected by PCR and two of the sequenced lines are described along with the trace files (Fig.6.1). Based on the sequence one *cipk-b* line has a 35bp deletion between the two sgRNAs resulting in a frameshift and the introduction of a premature stop codon resulting in a truncated protein product (1247-1). The second line (1247-2) for *cipk-b* has three areas of deletions, two of them are where the sgRNAs are targetted with one 5' upstream of the sgRNA1 site (Fig.6.1). The three areas of deletion are likely caused by two double strand break events followed by NHEJ events. Some homology can be found in areas between the cut sites. As the plant is haploid at this stage of the life cycle this cannot be caused by homologous recombinations and therefore must be caused two seperate NHEJ events, causing multiple mutations and deletions. Similarly 1247-2 results in a truncated protein product caused by a premature stop codon (Fig.6.1). Hence, two independent CRISPR/Cas9 mutants were generated and confirmed for *CIPK-B*. Growth of these two cipk-b mutants will be assessed in response to salt and zinc sulphate in order to determine if CIPK-B has a role in these stresses. Phenotypes from a single knockout of MpCIPK will also provide insight into whether the CIPKs are functionally redundant in *M. polymorpha*.

6.4.3 Cam2 and CIPK-B mutants do not respond to zinc sulphate up to 5 mM

Increasing zinc concentration did not result in a visible phenotype in Cam2 *M. polymorpha*. 5 plants were plated on 1/2 MS (1% sucrose) plates for 7 days with

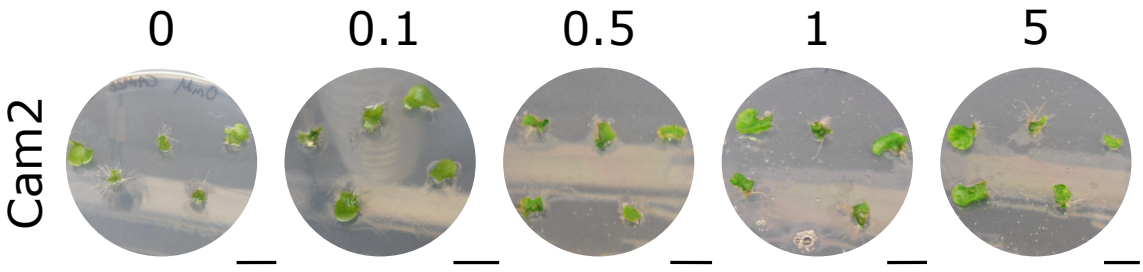


Figure 6.2: Response of Cam2 *M. polymorpha* to zinc sulphate. Wild type Cam2 *M. polymorpha* thallus cuttings were placed on 1/2 MS (1% sucrose) plates supplemented with zinc sulphate (0, 0.1, 0.5, 1, 5 mM) and show no visible phenotype after 7 day treatment. Black bars represent 1 cm scale.

increasing concentrations of zinc sulphate 0, 0.1, 0.5, 1, 5 mM per replicate for 5 biological replicates. No significant change in phenotype was visible after 7 days and plants were harvested for weight measurements in pooled samples containing 5 plants (Fig.6.2). Therefore it is unlikely that the dosage of $ZnSO_4$ that causes stress has been found for *M. polymorpha*. It may be that *M. polymorpha* can tolerate higher concentrations of $ZnSO_4$ for 7 days, or it may be that it takes longer for toxicity effects to accumulate. Therefore phenotypes in response to $ZnSO_4$ must be assessed over a longer time period, or at higher concentrations.

Increasing concentrations of zinc did not cause any change in fresh weight in the *cipk-b* lines or wild type Cam2 plants after 1 week. Fresh weight was measured for all replicates and found to be normally distributed and heterogenous. No significant difference was found for the increasing zinc sulphate concentration. However there was a significant difference in terms of growth in the *cipk-b* lines under all conditions compared to Cam2 including 0 mM zinc sulphate which would imply and overall reduction in growth on 1/2 MS (1% sucrose) (ANOVA: $P<0.01$). To control for the difference in terms of growth in the *cipk-b* lines compared to the wildtype each replicate was normalised to growth under 0 mM zinc sulphate but no significant difference was found in terms of growth on increasing zinc sulphate (Fig.6.3B). Therefore no detectable phenotype was found between Cam2 and both *cipk-b* lines after 7 days on concentrations of zinc sulphate up to 5 mM. However, as no phenotype was detectable for Cam2 on concentrations of zinc sulphate up to 5 mM after 7 days, it may be that a phenotype may be detected in *cipk-b* lines at

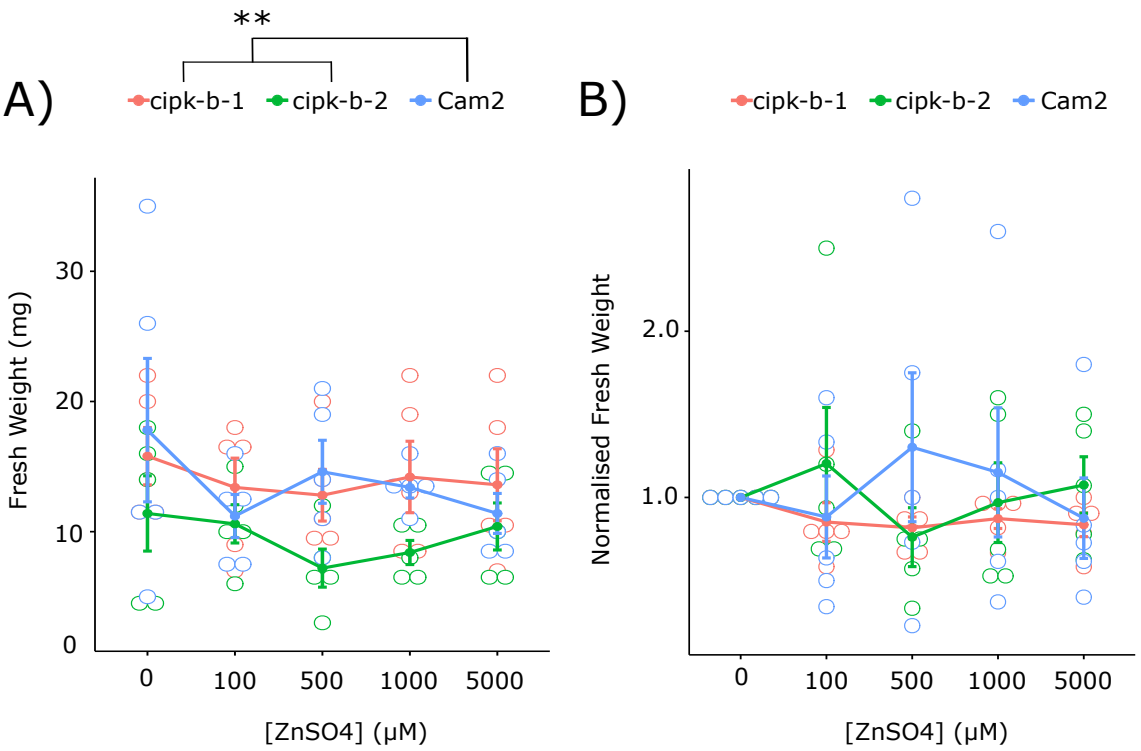


Figure 6.3: *M. polymorpha cipk-b* lines do not respond differently to zinc sulphate. *cipk-b* demonstrate no phenotype in response to zinc sulphate but have a reduced fresh weight and hence growth in all conditions including 0 mM. A) fresh weight of pooled plants with 5 plantlets per plate for two lines of CIPK-B mutants and Cam2 wildtype for 5 replicates after 7 days. B) Pooled plant weights normalised to the internal control of 0 mM zinc sulphate for all samples to access changes in freshweight caused by zinc sulphate. ** represents P<0.01, ANOVA.

higher concentrations of zinc sulphate or over a longer time period.

6.4.4 *cipk-b* knockout lines are more Salt Sensitive than Cam2

To access if the *cipk-b* lines operate in a salt tolerance pathway, *cipk-b* plants were compared to Cam2 for reduced growth on increasing salt concentrations. The two *cipk-b* lines were compared to wild type Cam2 and were found to be more sensitive to salt (Fig.6.4). The two lines of *cipk-b* can be seen to have a reduced growth from 50 mM NaCl compared to wild type Cam2 plants. Also *cipk-b* lines are visibly more chlorotic from 100 mM compared to Cam2 (Fig.6.4). Hence the *cipk-b* lines can be considered more sensitive to salt. However, growth can also be assessed by fresh weight to ensure this result is significant and not just due to reduced growth found

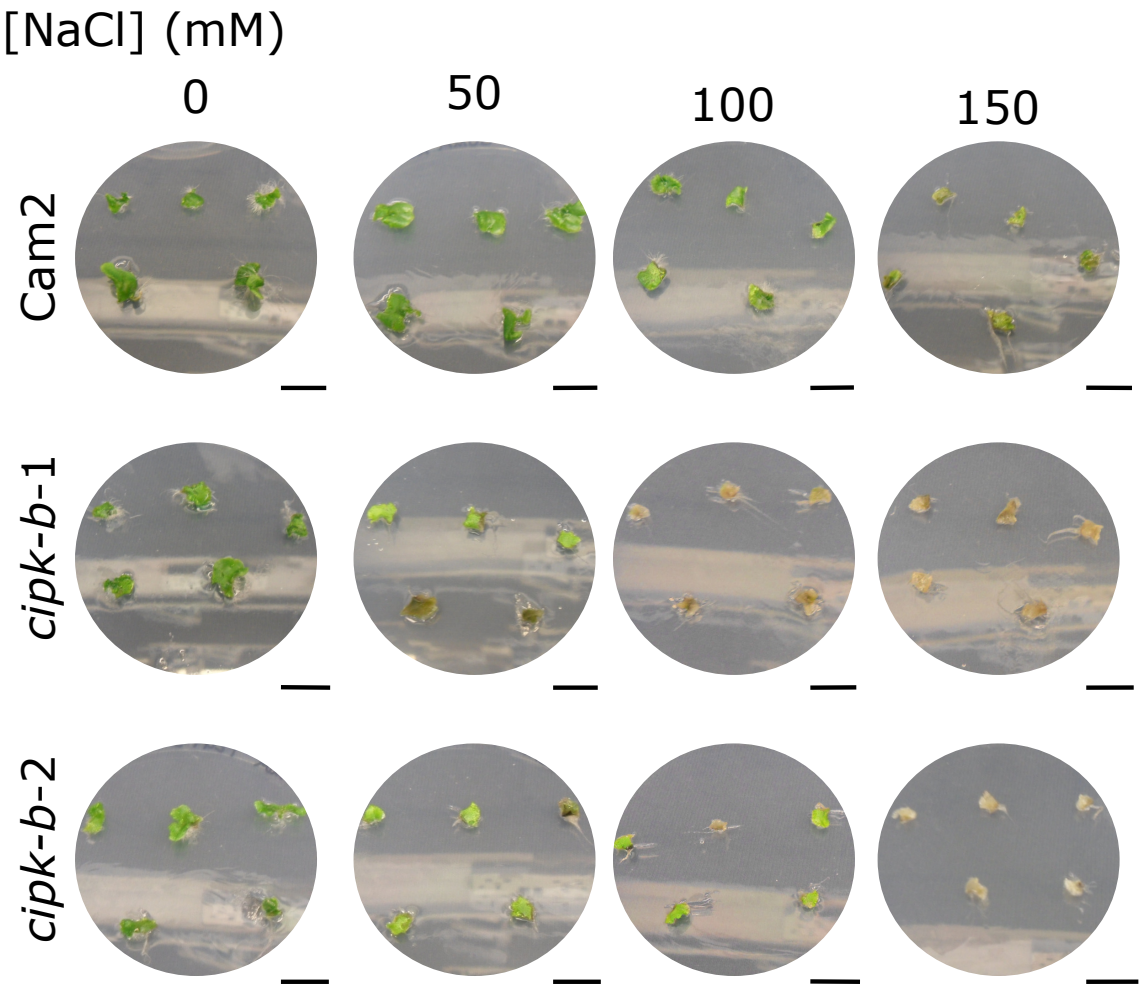


Figure 6.4: *M. polymorpha cipk-b* lines are visibly more sensitive to salt than Cam2. Wildtype Cam2 *M. polymorpha* thallus cuttings were placed on 1/2 MS (1% sucrose) plates with supplemented NaCl (0, 50, 100, 150 mM) for 7 days treatment in comparison to *cipk-b* lines. Black bars indicate 1 cm scale.

in the *cipk-b* lines.

the *cipk-b* lines had visibly reduced growth from 50 mM salt compared to with Cam2 plants. As *cipk-b* lines were shown to have significantly reduced growth in 0 mM zinc sulphate conditions (Fig.6.3A) it could be that the visibly reduced growth on salt is due to this growth defect (Fig.6.4; Fig.6.5.A). Therefore significance of the salt sensitivity was assessed in both fresh weight and when normalised to 0 mM salt to ensure overall difference in growth between *cipk-b* lines and Cam2 did not affect whether a phenotype could be detected. Fresh weight data showed the same reduced growth in *cipk-b* lines compared to Cam2 in 0 mM ($P<0.05$, Kruskall-Wallis). Fresh weight data also demonstrated reduced growth at 50 mM salt in both *cipk-b* lines

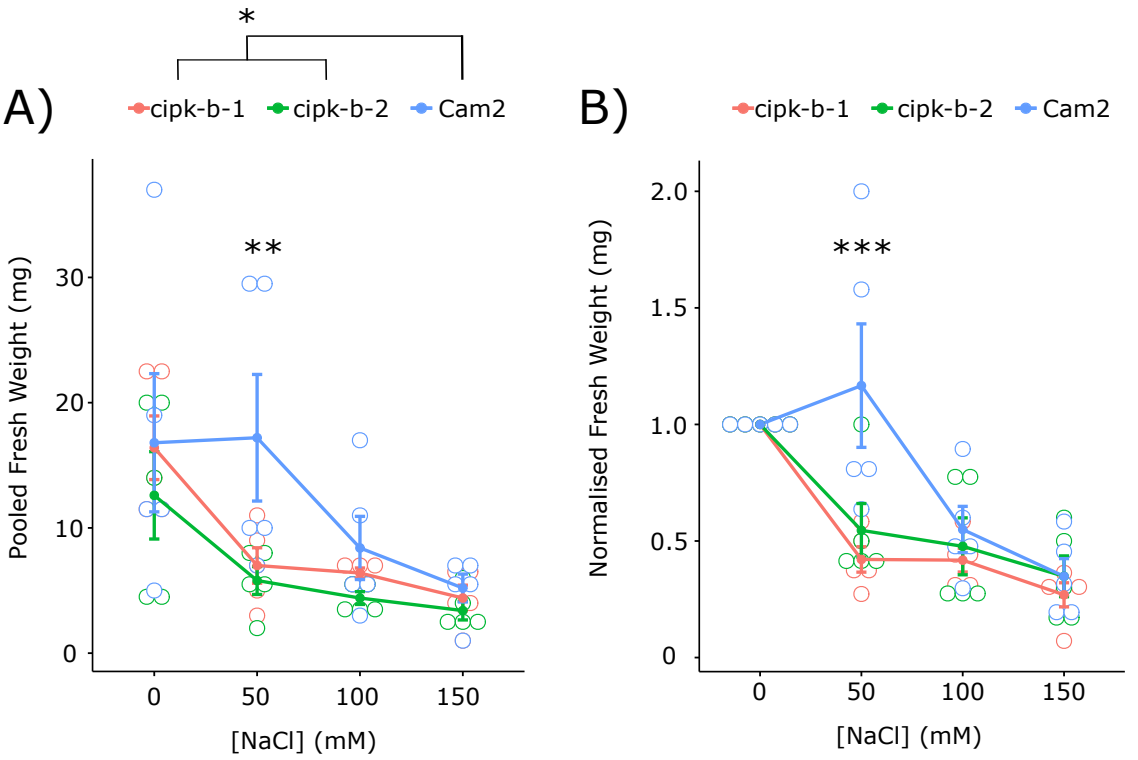


Figure 6.5: *M. polymorpha cipk-b* lines have reduced growth in response to salt stress. However *cipk-b* lines also have a reduced fresh weight and hence growth in all conditions. A) fresh weight of pooled plants with 5 plantlets per plate for two lines of CIPK-B mutants and Cam2 wildtype for 5 replicates after 7 days. B) Pooled plant weights normalised to the internal control of 0 mM NaCl for all samples to access changes in freshweight caused by NaCl. * = P<0.05, **= P<0.001, ***=P<0.0001.

compared to Cam2 (P<0.01, Kruskall-Wallis; Fig.6.5A). The reduced growth at 50 mM could not be detected at 100 mM or 150 mM salt due to the small amount of growth at higher salt concentrations. There was also no significant reduction in growth in 100 mM and 150 mM salt in Cam2 (Fig.6.5). While the data was found to be not normally distributed, or heterogenous, a significant difference was found between Cam2 and the *cipk-b* lines at 0 mM (Kruskal-Wallis: P>0.05). Therefore each replicate was normalised to the plants grown on 0 mM NaCl to control for changes in growth in the *cipk-b* lines, which also made the data heterogenous but not normally distributed. There was still a significant difference in growth between wild type and both *cipk-b* with increasing concentration of salt when normalised to 0 mM salt (Fig.6.5B; Kruskal-Wallis: P< 0.001). The main significant difference between Cam2 and the *cipk-b* lines was in growth at 50 mM (P<0.001, Kruskall-

Wallis). Hence, *cipk-b* lines are more sensitive to salt than Cam2, and therefore CIPK-B is involved in salt tolerance responses in *M. polymorpha*.

6.5 Discussion

6.5.1 Transformation of *M. polymorpha* and Knockouts generated

The CRISPR/Cas9 system utilised in this study is functional in *M. polymorpha*. However, the transformation protocols used and generated require some refinement to gain a higher efficiency. Neither the G-Agar trap nor the T-Agar Trap protocols could be replicated to generate mutants in our laboratory. Likely there are other factors that may affect transformation efficiency such as the *Agrobacterium tumefaciens* strain used to carry out the transformation as the G-Agar Trap protocol using strains GV2260, EHA101, EHA105, LBA4404 and MP90 had variable efficiencies (Tsuboyama et al., 2018). The GV3101 strain used in this study likely has a very poor transformation efficiency under the current methods. Tsuboyama et al., (2018) demonstrated that transformation efficiency could change when using micropore tape or parafilm during the co-cultivation period during transformation. The main explanation for this could be that humidity affects transformation efficiency with *Agrobacterium*-mediated transformation. However, general laboratory humidity would affect the starting point of the humidity within the plate, and therefore it may be that humidity is different in our laboratory which may have caused the different efficiency when using Agar Trap methods (Tsuboyama et al., 2018). These small unreported differences may make large changes in effective protocols, and hence reproducibility in transformation of *M. polymorpha*. Our modified sporeling transformation protocol had a relatively low efficiency of 0.3% but as 1000 gemmae can easily be used in the protocol this provides little difficulty in terms of generating mutants. In three replicates, of 1000 gemmae, per replicate we generated 2 CIPK-A CRISPR construct transformants in 1 replicate, 3 CIPK-B CRISPR construct

transformants in 1 replicate, and 4 CIPK-B CRISPR construct transformants in a third replicate. All of the produced transformants had the hygromycin casette or mutations in the genes of interest as detected by PCR and sequencing. This protocol could therefore be used to generate more transgenic lines but requires a relatively large amount of starting material.

The separate efficiency of CRISPR/Cas9 single guides could still provide difficulty in effectively producing multiple lines of mutants in *M. polymorpha* using the Cas9 system with our modified sporeling transformation method. Of the 7 CIPK-B lines that were generated only two had detectable mutations by PCR and both of those lines introduced premature stop codons and hence generated knockout mutants. Other lines of CIPK-B and CIPK-A are still being assessed in the laboratory for the nature of any caused mutations (Personal Communications, Althea Rose). However with a low transformation efficiency, and a low detection of mutations caused by two sgRNA target cleavages by PCR, it may be that the single guides used do not have the highest efficiency. Therefore differential efficiency of guides may make future gene deletion lines difficult to generate if efficiency of single guides is lower in other CRISPR/Cas9 constructs. Although only a single sgRNA needs to target the Cas9 enzyme to cause a double strand break for a possible deletion mutant line in the gene of interest therefore more mutants may be possible to generate but these mutants may not be detectable by PCR.

6.5.2 *cipk-b* lines may not be more Sensitive to Zinc Sulphate

Cam2 was not sensitive to the zinc sulphate up to 5 mM after 7 days and the *cipk-b* lines also showed no sensitivity. However, *cipk-b* lines do seem to have reduced growth on 1/2 MS (1% sucrose) media compared to wildtype Cam2 plants. IRT1 is known to be phosphorylated for degradation when heavy metals, such as zinc and cadmium, exceed iron concentration. As the IRT1 protein lacks specificity towards iron uptake and, in those conditions, other heavy metals can be taken up which can

be toxic (Dubeaux et al. 2018). In *A. thaliana*, AtCIPK23 is responsible for the phosphorylation of IRT1 which allows it to then be degraded by a ubiquitin degradation pathway (Dubeaux et al. 2018). Of the known CIPK functions, heavy metal toxicity has no overlap with salt stress as it does not relate to osmotic stress or ionic balancing caused by sodium stress, which was why zinc tolerance was tested. The *cipk-b* lines demonstrated no distinct reduction in growth or plate visible stress after 7 days at 5 mM zinc sulphate, but neither did the wild type control, and therefore it can be difficult to ascertain whether a phenotype would be expected (Fig.6.2 , Fig.6.3). It may be that a stress response would have been demonstratable at a higher concentration of zinc sulphate or over a longer time-frame. However it is interesting that in *A. thaliana* adult plants only 23.4 % can survive in 0.5 mM $ZnSO_4$ for 15 days (Xu et al. 2010a). Similarly *N. tabacum* two week old plants show perturbed growth from 0.25 mM $ZnSO_4$ after 5 days (Vera-Estrella et al. 2017). Therefore compared to a number of angiosperm plants, *M. polymorpha* is much more tolerant to $ZnSO_4$ which concurs with the finding that *M. polymorpha* can survive 2 mM $ZnSO_4$ for three weeks (Ares et al. 2018). However other species of metalophyte have much higher tolerances such as *Arabidopsis halleri* which can survive 30-fold the zinc concentration of *A. thaliana* (Macnair et al. 1999). Therefore *M. polymorpha* can not be considered a true metalophyte but demonstrates surprisingly high tolerance which may be useful in investigating metal tolerance characteristics.

6.5.3 CIPK-B is involved in Salt or Drought Tolerance

The *cipk-b* lines had significantly reduced growth compared to Cam2 plants and also a reduced salt tolerance compared to wild type plants. The *cipk-b* lines were visibly more sensitive to salt from 50 mM NaCl with reduced growth and increased chlorosis (Fig.6.4). Fresh weight of *cipk-b* lines was also reduced after a when plants were grown under salt stress demonstrating a very low tolerance to salt compared to wild type plants (Fig.6.5). The *sos2* lines generated in *A. thaliana* also had reduced growth at 50 mM NaCl (Zhu et al. 1998). Therefore this may be a good

indication that CIPK-B is a SOS2 homologue instead of CIPK-A. Other phenotypes were also tested in *sos2* including a sensitivity to low K^+ conditions which could be tested in the *cipk-b* lines. Although several other AtCIPKs are also sensitive to low K^+, including AtCIPK23 (Cheong et al. 2007; Ragel et al. 2015). Equally the reduced growth under salt stress in *cipk-b* lines could actually be a reduced tolerance to osmotic stress which is caused in salt conditions. Osmotic stress could be determined by assessing growth of *cipk-b* lines on sorbitol or mannitol for an equal osmotic stress (i.e. 100 mM to 50 mM NaCl) to gain resolution on if the stress susceptibility is cause by the ionic factor of NaCl. As omotic stress susceptibility is not a phenotype of the *sos2* lines in *A. thaliana cipk-B* lines could be tested for osmotic stress susceptibility. If no osmotic stress susceptibility is found this could provide more evidence that CIPK-B is a SOS2 homologue. Similarly it has been previously shown that the SOS2 mutant (*sos2*) shows growth reduction compared to wildtype on 1/2 MS plates which can be recovered by supplementing 50 mM KCl (Zhu et al. 1998). Therefore it may be possible to counter the effects of reduced growth in *M. polymorpha cipk-b* lines by supplementing 50 mM KCl. Overall the sensitivity to salt is an important phenotype detected in *cipk-b* lines, however more testing is required to completely characterise the phenotypes of these mutants. Some indications imply CIPK-B could be a SOS2 homologue present in *M. polymorpha* but further phenotypic testing and characterisation of a downstream Na^+/H^+ antiporter would be required to confirm this.

A number of key experiments would be required to completely confirm that CIPK-B is a SOS2 homologue and that the SOS pathway is conserved between *A. thaliana* and *M. polymorpha*. The most important experiment to confirm this would be to assess if CIPK-B can complement *sos2-1 A. thaliana* lines. To do this CIPK-B would be placed under a constitutive promoter and transformed into *sos2-1* and salt sensitivity can be assessed compared to wildtype *A. thaliana* and *sos2-1* to determine if salt sensitivity can be reduced compared to *sos2-1*. A similar experiment was carried out in *P. patens* to determine that PpCIPK1 is a SOS2 homologue (Xiao

et al. 2021). Alongside this bimolecular fluorescence complementation can be carried out for CIPK-A and CIPK-B to determine if an interaction is possible between AtSOS1 and CIPK-B from *A. thaliana*. Alongside this the SOS1 homologue(s) can be determined bioinformatically similar to how the CIPKs were identified in *M. polymorpha* (Chapter 3) and BiFC can be used to check if interactions are possible between CIPK-B and the SOS1 homologue(s) in *M. polymorpha*. If these experiments can determine that CIPK-B can complement *sos2-1 A. thaliana*, the MpSOS1 homologue(s) can be determined, and interaction can be confirmed then CIPK-B can be confirmed as a SOS2 homologue present in *M. polymorpha*.

In conclusion, it is tempting to suppose that CIPK-B interacts with our potential SOS3 homologue CBL-C in *M. polymorpha* and constitutes a SOS2 homologue but further experimentation is required to determine this conclusively. Most importantly it would be key to separate out the potential of *cipk-b* lines being osmotic stress susceptible, or a salt ionic stress susceptible. Also it would be important to determine the downstream Na^+/H^+ antiporter that the CIPK may be phosphorylating as a SOS1 homologue to truely determine if this is a conserved pathway between *A. thaliana* and *M. polymorpha*. The detection of a phenotype from *cipk-b* lines does provide evidence of a lack of complete functional redundancy in *M. polymorpha* CIPKs which is a key finding that encourages further work in this system. It may be that CIPK-A also works in salt or drought susceptibility and is somewhat redundant but the detectable phenotype can at least be used to confirm function of *cipk-b* in salt. Knockout *cipk-a* lines still require phenotyping to conclusively determine a lack of functional redundancy between the CIPKs in *M. polymorpha*.

Chapter 7

Discussion

7.1 CIPK-B is involved in salt or drought responses and may indicate evolutionary history of the CBL-CIPK pathways

The CBL-CIPK pathway has massively expanded and diversified with the emergence onto land and then subsequent expansion of the green lineage. Of the pre-terrestrial plants, *Ostrococcus tauri* is a chlorophyte with a single CBL and CIPK, *Klebsormidium flaccidum* has a single CIPK and 3 CBLs, therefore the toolkit was present prior to pre-terrestrialisation (Kleist et al. 2014). The bryophytes have a complex phylogeny which has recently been resolved into a monophyletic group firstly with liverworts and mosses (Puttick et al. 2018). Subsequent studies, including wider sampling of bryophytes, trait analysis, and sequencing of the genome of a hornwort (*Anthoceros angustus*) has consistently provided more evidence for the monophyly of bryophytes with hornworts as a sister clade to setaphytes (mosses and liverworts) (Zhang et al. 2020; Su et al. 2021; Harris et al. 2020). Much of the difficuty in studying the evolution and diversification of signalling families is linked to the difficulty of resolving early land plant evolution. Bryophytes are the second most diverse group of land plants (Laenen et al. 2014) which has made it difficult to understand which traits are ancient and are caused by the expansion and diversity of the group.

The poor fossil history (Tomescu et al. 2018) of bryophytes has added difficulty to ascribing ancient traits and time resolving their evolution. A range of historical expansions and extinctions have been reported for bryophytes (Laenen et al. 2014) which has made understanding of the key traits in signalling, stucture and defence that led to the emergence onto land difficult to parse. The bryophytes have continued to evolve with the loss of important traits that were key to the terrestrialisation events, including the loss of stomata and guard-cell function in liverworts as an example (Harris et al. 2020). Therefore the only way to effectively understand the evolutionary origin of signalling components that expanded with emergence onto land is to use a macroevolutionary process of analysis of a number of bryophytes and comparing traits across these groups.

To carry out macroevolutionary research on CBL-CIPK expansion and original function, functions will have to be ascribed to CBLs and CIPKs in bryophytes. Work in this thesis only found 3 CBLs and 2 CIPKs in *M. polymorpha* which confirmed the results of Edel and Kudla (2015) which used version 3 of the genome, so no new CBLs or CIPKs have been detected in version 5 of the genome. Other work has characterised 7 CIPKs and 4 CBLs in *P. patens* (Kleist et al. 2014). *A. angustus* has not been previously analysed and the genome is not currently of a high enough quality (Zhang et al. 2020) to assess the number of CBLs and CIPKs but at least 3 scaffolds seem to harbour CIPK sequences upon initial inspection with tBLASTn using AtCIPK sequences (VJWM01000172.1, VJWM01000159.1 and MU008195.1). However, full sequences could not be retrieved for these CIPKs (Data not shown). Higher quality sequence data would be required to retrieve these sequences and molecular biology tools need to be developed to fully interrogate the system and function of these CIPKs in *A. angustus*. Initial work has been done on characterising CIPK-B in *M. polymorpha* and its role in drought or salt responses in this thesis. However work in *P. patens* has also demonstrated the salt sensitivity response with the *PpCIPK1* knockout (Xiao et al. 2021). Therefore we can at least suspect that drought and salt responses are highly conserved from bryophytes to angiosperms.

Confirmation of a similar phenotype in *A. angustus* would be helpful to demonstrate that all clades of bryophytes have a similar response. However the analysis across *M. polymorpha* and *P. patens* of CIPK sequences demonstrates expansion of CIPKs in *P. patens* which may be linked to its whole genome duplication events providing evolutionary reasoning for expansion. *A. angustus* has no evidence of genome duplication and therefore it is likely to have a similar number of CBLs and CIPKs and therefore further research in *A. angustus* would be critical to make broad assumptions about the evolutions of these networks as well as original functions.

CDPK	3	7	21	10	13	34	29	41
CBL	1	3	4	4	7	10	11	12
CIPK	1	2	7	5	26	26	34	43

Figure 7.1: Phylogeny for the green lineage with numbers of CDPKs, CBLs and CIPKs. Phylogeny is proposed based on Puttick et al. (2018) and Harris et al. (2020). Numbers of CDPKs, CBLs and CIPKs are sourced from Edel and Kudla (2015), Kleist et al. (2014), Weckwerth et al. (2014) and Chapter 3.

Emergence onto land required development of multiple types of tolerance to abiotic stress compared to the aquatic green lineage. Early land plants had to develop an ability to handle UV radiation and other DNA damaging agents which has been suggested to have taken place in streptophyta through horizontal gene transfer from

bacteria (Fang et al. 2017). Other stresses include dessication and rehydration, temperature changes, metal stress and a higher concentration of carbon dioxide. Green algae species only have single gene pairs of CBLs and CIPKs including: *Coleochaete orbicularis, klebsormidium flaccidum, Chaetosperidium globosum* and *Penium margaritaceum* (Kleist et al. 2014). Whereas *M. polymorpha* and *P. patens* have 3 and 4 CBLs, and 2 and 7 CIPKs, respectively (Fig.7.1). The number of CBLs and CIPKs demonstrate a clear increase in the number of CBLs and CIPKs from green algae to bryophytes including in *M.polymorpha* which has no evidence for whole genome duplication (Bowman et al. 2017). The defined functions of CBLs and CIPKs include ABA responses including drought, potassium deficiency, high salinity and toxic metal tolerance (Beckmann et al. 2016). Importantly CIPK-B has been defined as being involved in drought or salt stress in *M. polymorpha* (Fig.6.5) and a single CIPK has also been defined in *P. patens* as salt tolerance conferring. All of these defined functions cover the stresses that are recognised for the ability of plants to emerge onto land and therefore CBLs and CIPKs may represent a core signalling component that expanded for plants to be able to emerge onto land.

As these CBL CIPK components expanded with the emergence onto land, whole genome duplication (WGD) also plays a large role in the expansion and allow for neofunctionalisation. Analysis of the CBLs and CIPKs in *P. patens* has found that most of the CBLs and CIPKs form paired cognate loci including PpCBL2/3, PpCIPK1/5, PpCIPK3/4, PpCIPK6/7 with unpaired PpCBL1, PpCBL4 and PpCIPK2 (Kleist et al. 2014). A number of these PpCBL genes seems to be in various levels of psuedogenisation including PpCBL6 which does not seems to be expressed in any known conditions. Gene loci could not be investigated in *M. polymorpha* as the genome is not currently physically mapped. However, all of the CBL and CIPK genes in *M. polymorpha* are expressed, at least as mRNA, which was detectable through qRT-PCR and RNA sequencing (Chapter 4). Alongside this, a phenotype was detectable from a single gene knockout of *cipk-b* which implies that the CIPKs are not functionally redundant. Different transcriptional responses between *CBL-B*

and *CBL-C* during salt stress (Fig.4.8), and likely different localisation for CBL-A (Fig.3.2) implies that functional redundancy is also unlikely between CBLs in *M. polymorpha*. The pairing post-WGD likely causes much of the observable functional redundancy found in these protein families in later-diverging plants (Kleist et al. 2014). As the WGD event likely caused the formation of cognate pairs in *P. patens* and rapid expansion of the calcium decoder network it also allowed for psuedogenisation to take place due to reduced selection to keep functionally redundant genes. This reduced selection also allows for neofunctionalisation as selection pressure is removed and mutations can allows for the development of new functions for the network. Therefore whole genome duplication and likely hybridisation events can allow for new functions to be developed in the calcium decoder network in angiosperms. Therefore the only way to determine what the original functions of the CBL and CIPK network were, is to carry out comparative investigation into CBL and CIPK functions across plants that diverged near the emergence onto land, including multiple groups of bryophytes to complement the work being carried out on angiosperms.

7.2 Salt-induced Calcium Signalling needs further Elucidation

Analysis of the calcium signals triggered by salt and drought stress (sorbitol) have demonstrated different signals in response to each stress in *M. polymorpha*. The differences between signals caused by drought and salt are also true in other land plants such as *A. thaliana* (Schmöckel et al. 2015). The signals generated in response to salt and drought take place in a shorter time frame in *A. thaliana* than in *M. polymorpha* but in *M. polymorpha* salt signals are still characteristically biphasic and drought signals are monophasic on a whole-cell level (Schmöckel et al. 2015; Huang et al. 2017). In *A. thaliana* it was determined that the biphasic signal in response to salt is actually two monophasic signals localised to different subcellular locations (the nucleus and the plasma membrane) that are independent. The monophasic

signal in *A. thaliana* in response to osmotic stress was also two monophasic signals that overlap and are localised to the nucleus and plasma membrane (Schmöckel et al. 2015). The same resolution of signal could not be carried out with the Tak1 R-GECO line as the construct was not targeted to any subcellular location (Fig.4.1). Both osmotic and salt stress induced calcium signals are independent of each other in *A. thaliana* as salt stress induced nuclear and cytosolic calcium release can inhibit root primary growth and osmotic stress induced calcium signals cannot (Schmöckel et al. 2015). However the independence of the two stress signals is not known in *M. polymorpha* and it would be interesting to determine if the osmotic stress induced calcium signal is a component of salt stress induced calcium signal, as osmotic stress is a component of salt stress. The signal could be looked at with greater resolution *M. polymorpha* to determine if the two parts of the biphasic signal take place in different subcellular compartment. This could be carried out by targetting fluorescent reporters of calcium signals to subcellular compartments as has been previously done (Huang et al. 2017). If the salt induced biphasic signal is actually two monophasic signals at different subcellular compartments that overlap on a whole cell level, fluorescent reporters of calcium signals targeted to different subcellular compartments can be used to interrogate if these signals may also be drought induced. Downregulation of calcium decoding proteins may imply that the signals are continuous over time but the signals have not been previously checked to see if they happen continuously over a week of salt stress, or if they eventually abate and with which frequency they continue to signal if they do.

Alongside the duration of the signal, the distance of the signal should also be considered. Long distance calcium signals have been recently published for a variety of stress in plants including in response to glutamate and wounding responses associated with herbivory (Shao et al. 2020). Root-to-shoot calcium signals have been defined for salt stress in *A. thaliana* travelling through cortical and endodermal cells at 400 μm/s (Choi et al. 2014). Alongside this a mannitol triggered drought stress signal was also assessed and demonstrated that the signal stays predominantly in

the roots with a very weak signal in the leaf in *A. thaliana* (Zhu et al. 2014). These distance of the signals and the transmission would be very interesting to investigate in *M. polymorpha* due to the plants different structure with unicellular rhizoids and thallus instead of traditional roots and shoot. Also the signal in *M. polymorpha* is much slower and it would be interesting to see if this translates into a rate difference if *M. polymorpha* is shown to also have systemic signals. Also it was found that over-expression of TPC1 increases the rate of transmission in *A. thaliana* and increases expression of salt tolerance genes. It would therefore be interesting to investigate these responses in *M. polymorpha* due to the slow rate of the salt induced calcium signal. Interestingly TPC1 is localised to the tonoplast and yet seems to coordinate the long distance signalling in response to salt and yet the intracellular salt signalling has been defined as nucleus and plasma membrane localised. Alongside this long term signalling in response to salt has not been defined and yet there is some circumstantial evidence that signalling continues after the initial salt stress. Much more work is required to localise signal detection to subcellular compartments and investigate short and long term responses to salt. Knockout of different channels including GLR3.7 at the nucleus, and TPC1 at the tonoplast, is also required to separate out which calcium decoders are activated at different subcellular localisations to understand the functions of different calcium decoders in response to salt.

7.3 Transcriptional control of Calcium decoding proteins implies roles in Early and Late Salt tolerance

Studies in CDPKs seem imply that the salt and drought responsive CDPKs are involved in continuous long term signalling in response to these stresses. Many of the previous studies on CDPKs have focussed on expression of genes in the first 24 hours after salt treatment for a range of plants and different concentrations of salt. AtCPK3 has been demonstrated to be upregulated within an hour (Mehlmer et al.

2010). AtCPK6 has been shown to be upregulated from 2 hours (Xu et al. 2010b). AtCPK10 is upregulated after 2 hours of salt treatments (Zou et al. 2010). Studies in other plants have also taken the same approach to define salt and drought responsive CDPKs including in *Nitraria tangutorum, Fragaria x ananassa* and *Brachypodium distachyon* (Crizel et al. 2020; Wen et al. 2020). However other studies have show that negative regulators of salt and drought response can be upregulated initially and are downregulated at later time points (Gong et al. 2005; Li et al. 2016). Therefore studies may focus on trying to identify salt and drought tolerance genes and then discover they are actually negative regulators. In *M. polymorpha CDPK-C/E/F/G* are upregulated from 0-24 hours (Fig.4.10) and *CDPK-E/G* could be detected as upregulated after 1 week 50 mM treatment (Fig.4.9) which implies they are both involved in long term continuous salt responses. *CDPK-E/F* could also be detected by qRTPCR at higher concentrations of salt from 100-150 mM after a week (Fig.4.6) and therefore it is also involved in long term responses at higher concentrations of salt. Hence *CDPK-E* is the best candidate for conferring salt tolerance of the CDPKs as it is upregulated in the short term (0-24 hours) and long term (7 days) at low (50 mM) and higher (100-150 mM) concentrations of salt (Fig.4.10; Fig.4.9; Fig.4.6). As phenotyping of plants in response to abiotic stresses typically takes place after a week, and due to the confounding nature of early upregulation of CDPK negative regulators of salt and drought responses, future studies should consider checking for expression changes in candidate genes at longer term time points to look for consistent results.

As *CDPK-E/F* are both induced by salt and as they are still upregulated after 7 days they are likely involved in tolerance mechanisms rather than as negative regulators. The salt tolerance induced CPKs in *A. thaliana* include responses controlling H_2O_2 production (AtCPK27) (Zhao et al. 2015), ABA-dependent salt responses (AtCPK6) (Xu et al. 2010b) and stomatal regulation (AtCPK3) (Mehlmer et al. 2010). Of these responses AtCPK3/27 are in subgroup II and CPK6 is in subgroup I. Neither CDPK-E/F are in these subgroups and therefore it may be likely that

these CDPKs are actually responsive to drought conditions caused by salt stress. CDPK-F is in subgroup III and is therefore more similar to AtCPK10 which is involved in drought responses through stomatal regulation (Zou et al. 2010). Also in subgroup III is AtCPK13 which is light induced and inhibits stomatal opening in response to drought stress. AtCPK8, from subgroup III, is involved in H_2O_2 and stomatal regulation with interaction CAT3 (Harris et al. 2020). Therefore the upregulation of *CDPK-F* is surprising as stomata are not present in *M. polymorpha* and there is no evidence for changes in conductance of air pores either. The only other known AtCPKs involved in salt induced calcium signals are AtCPK6/16/34 which are in subgroup I/IV/III respectively and have been shown to phosphorylate GLR3.7 in *A. thaliana* to increase cytosolic calcium release and salt tolerance (Wang et al. 2019). As there is no known change in air pore conductance and stomata do not exist in *M. polymorpha* it may be likely that CDPK-F is drought responsive but through a novel mechanism. CDPK-E may function in a similar manner to AtCPK16 which is supported by the EF hand analysis (Chapter 3.2.3) and therefore it may be involved in maintainence of salt induced calcium signals. However the EF hand analysis of CDPK-F implied closer EF hand identity with AtCPK13/30 therefore there could still be a novel mechanism for CDPK-F activity. As *CDPK-F* seems to be upregulated by salt and it is one of 20 genes that have homologues on both chromosome U and V in *M. polymorpha* (Bowman et al. 2017) there seems to be selection for it to be present in both sexes of *M. polymorpha*. As tolerance to salt responses seem consistent between the sexes in both Cam and Tak accessions it is difficult to currently propose a function for CDPK-F but it would be a very interesting CDPK to investigate further.

CBLs seem to activate predominantly in short term salt responses including those from *M. polymorpha* and *A. thaliana*. The Salt Overly Sensitive pathway is upregulated quickly in the root in response to salt with increase in *SOS3* and *SOS1* expression after 1 hour in the endodermis and epidermis respectively (Ji et al. 2013). *SCaBP8 (AtCBL10)* increases in expression slightly later at 24 hours likely as the

root SOS pathway mitigates most of the salt entering the shoot until this later time point (Quan et al. 2007). All of the components of the SOS pathway are downregulated after 6 days implying the SOS pathway is an early salt responsive pathway in *A. thaliana* (Rolly et al. 2020). *CBL-A/B/C* are all upregulated in the first 24 hours in response to 100 mM NaCl (Fig.4.10). However only *CBL-C* is downregulated after a week in RNA sequencing and qRTPCR in Tak1 *M. polymorpha* (Fig.4.9; Fig.4.6). It is tempting to propose that *CBL-C* is a SOS3 homologue in *M. polymorpha* as it is also an early responder to salt stress and downregulated at 7 days (Fig.4.8). Tissue localisation of these proteins would be incredibly interesting to investigate as *M. polymorpha* only has unicellular rhizoids instead of multicellular roots and investigation of this may help to understand how the complexity of the SOS system developed from bryophytes to angiosperms. It may be interesting if CBL-C is only found in rhizoids as these cells are more similar to root hairs than true roots and yet the tissue localisation would imply that the rhizoids create a barrier to salt stress affecting the rest of the plant, similar to how different SOS pathways are present in the roots and shoots of *A. thaliana*.

CIPK-A and *CIPK-B* are both upregulated after 24 hours and downregulated after 1 week in Tak1 and Cam2 (Fig.4.10; Fig.4.9; Fig.4.8; Fig.4.7). However, after a week, Tak1 *CIPK-A* was only downregulated at 50 mM whereas *CIPK-B* was downregulated at 50-100 mM and in Cam2 it is downregulated in 100-150 mM NaCl (Fig.4.7). *CIPK-B*, on the otherhand, was downregulated at 50-100 mM in Tak1 and 50-150 mM in Cam2 (Fig.4.7; Fig.4.8). However by RNA sequencing *CIPK-B* was upregulated after a week (Fig.4.9). Due to the inconsistency in these results caution was applied to ascribing function. However the *cipk-b* lines demonstrated increased sensitivity to salt it is likely that CIPK-B is a SOS2 homologue or at least operates in salt and drought responses. However *SOS2* is also not known to be regulated in expression in *A. thaliana* which may be why the changes in expression of *CIPK-B* were inconsistent. Perhaps there may be more regulation of the SOS pathways through the regulation of expression of *SOS3*. The lack of a consistent

change of expression of *CIPK-B* may actually provide some indication that it is a SOS2 homologue or at least that the pathway is regulated in the same way. However it is intersting that there do seem to be some form of transcriptional changes in *M. polymorpha*. Potentially CIPK-B is regulated in response to other stresses or conditions not yet studied. Noticeably *CIPK-B* was upregulated in Tanaka et al. (2018) which could be due to different lifephases of *M. polymorpha* tested (Fig.4.9), but no developmental differences have been noticed for *cipk-b* lines beyond reduced growth (Fig.6.5). However if CIPK-B is a SOS2 homologue this may be recoverable in increased KCl as was found in the *sos2-1*. Therefore potentially the difference between Tanaka et al. (2018) and the qRT-PCR could be the media used and as CBLs and CIPKs are known to be involved in ionic control it is likely that the different ionic compositions of the media affected expression.

The Salt Overly Sensitive pathway is a conserved pathway from bryophytes such as *Physcomitrium patens* to angiosperms such as *A. thaliana* (Xiao et al. 2021). The demonstration of similar transcriptional control between *CBL-C* and *SOS3* alongside the salt sensitivity phenotype of *cipk-b* in *M. polymorpha* and interaction between CBL-C and CIPK-B make a persuasive argument for the existence of a Salt Overly Sensitive pathway in *M. polymorpha* (Fig. **??**. Similarly in *P. patens* SOS1 homologues have been defined with two different isoforms capable of mediating sodium efflux (Fraile-Escanciano et al. 2010). A CIPK (PpCIPK1) has also been defined as a SOS2 homologue in *P. patens* and can functionally complement *sos2-1* knockout line in *A. thaliana* (Xiao et al. 2021). One aspect that is lacking in both of these bryophyte systems is the tissue specificity of the Salt Overly Sensitive pathway. In *A. thaliana* it is proposed that *SOS3* is upregulated in the endodermis to modulate and increase endogenous SOS1 activity in response to salt (Fig.7.2). SOS1 is upregulated in the epidermis of roots if the increased activity in the endodermis is overcome by large amounts of salt (Fig.7.2). There is evidence that SOS3 overexpression causes the same increase in salt tolerance as overexpressing SOS1 (Yang et al. 2009). *AtCBL10* expression is increased in the shoot after 24 hours if

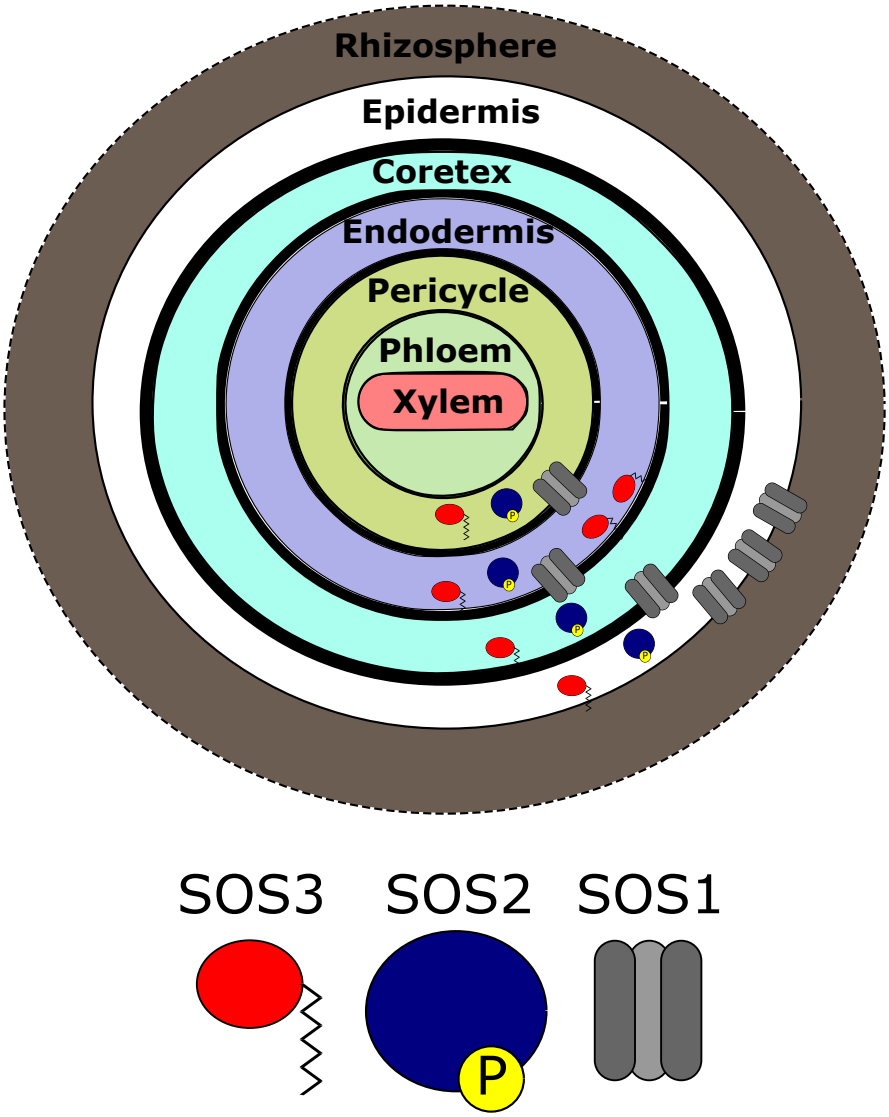


Figure 7.2: Model of the SOS pathway in *A. thaliana* root cross-section. The layers of *A. thaliana* root cell types are labelled within the cell layer. The SOS pathway is expressed in all parts of the root but expression of *SOS1* is higher in the epidermis represented here as presumably higher number of SOS1 protein channels. *SOS3* has higher expression in the endodermis represented here by presumably more SOS3 protein. The increased expression likely leads to higher export of sodium from these tissues and form part of the defence against salt entering the transpirational flow (xylem).

the traditional SOS pathway is overcome and too much sodium enters the transpiration flow and risks damaging delicate floral tissues. If CBL-C and CIPK-B, and PpCIPK1 and PpSOS1, present parts of SOS pathways in bryophytes it would be important to define the tissue localisation that form part of this pathway to make interesting macroevolutionary comparison to understand different mechanisms that can confer salt tolerance in higher plants.

7.4 Multiple interaction test are required to define CBL-CIPK interactions

Yeast two-hybrid has been a staple of assessing interactions between CBLs and CIPKs but may be a poor method to define interactions. Yeast two-hybrid has been used to test interactions between CBLs and CIPKs from various plants fusing the GAL4 binding domain to the CBL N-terminus and the GAL4 activation domain to the CIPK N-terminus and preceeded to test interactions (Halfter et al. 2000; Albrecht et al. 2001; Hashimoto et al. 2012). Yeast two-hybrid have shown different resulting interactions when reportedly using the same constructs including reporting an interaction between AtCBL1 and CIPK1 (Kolukisaoglu et al. 2004) or no interaction for the same proteins (Albrecht et al. 2001), an interaction between AtCBL1 and AtCIPK7 (Kolukisaoglu et al. 2004) or no interaction between these proteins (Albrecht et al. 2001) and, a weak interaction between CBL1 and CIPK12 (Kolukisaoglu et al. 2004) or no interaction (Albrecht et al. 2001). It is rare for yeast two-hybrid results to be re-reported so it is difficult to know how reproducible many of the results are from different labratories. Alongside this, in original papers many of the interactions are defined by just yeast two-hybrid and not confirmed with other interaction tests. Some papers now analyse only a subset of interactions that were tested with yeast two-hybrid using other methods such as bimolecular fluorescence complementation (Zhang et al. 2014). The yeast two-hybrid to test for interactions in *M. polymorpha* found that CBL-C and AtCBL-4 were incapable of interacting with CIPK-A in yeast two-hybrid (Fig.5.1). Yet the bimolecular fluorescence complementation found that all interactions were possible (Fig.5.3; Fig.5.4; Fig.5.5) which demonstrated the need to further test many of the interactions reported in higher plants as the CBL-CIPK pathway may be hiding unique interactions that have not been thus far detectable.

Yeast two-hybrid has already proven unreliable in representing *in planta* CBL-CIPK interactions. AtCBL9 and AtCIPK3 were shown to not interact by yeast

two-hybrid (Kolukisaoglu et al. 2004) and yet they do interact in regulating ABA signalling in *A. thaliana* (Pandey et al. 2008). AtCBL4 was demonstrated to not interact with AtCIPK6 (Albrecht et al. 2001) but does, and is involved in mobilising AKT2 from the endoplasmic reticulum to the plasma membrane (Held et al. 2011). AtCBL3 was also shown not to interact with AtCIPK9 (Albrecht et al. 2001) however *in planta* this pathway is involved in K+ distribution and translocation (Liu et al. 2013). Therefore yeast two-hybrid can be a good initial test to find many interactions but all of those interactions have to be confirmed with further tests due to false positives and false negatives. The reason for this may be that there are multiple ways in which CBL-CIPK interactions may be facillitated which would not be represented in non-plant heterologous systems.

Requirements of calcium occupancy, phosphorylation and posttranslational modifications for interaction may cause difficulty in determining interactions between CBLs and CIPKs more generally. Many interaction tests use heterologous systems such as yeast two-hybrid or bimolecular fluorescence complementation to determine interactions and they have benefits and issues depending on the system that the protein is expressed in. Proteins expressed from heterologous systems can have different basal calcium concentrations, different posttranslational modification and may or may not be phosphorylated each of which could affect interactions between CBLs and CIPKs. BiFC can use plant systems that are likely more similar to other plant systems and therefore is more likely to be similar in terms of basal calcium and posttranslational modifications. Other interaction tests use *in vitro* methods such as isothermal titration calorimetry (ITC) which uses purified protein and titrates the suspected interactor and can be used to determine the interaction affinity for different CBLs and CIPKs to see if there is any preference in terms of interaction affinity. However ITC uses proteins expressed heterologously but they also can have unreliable results as generally interactions are assesed in buffers that do not represent *in vivo* conditions. Co-IP is perhaps one of the best current systems for testing interactions for CBLs and CIPKs in higher plants as the test uses natively

expressed proteins under native promoters. However Co-IP can be laborous processes of adding tags to proteins of interest for later purification and expressing it in the native plants and the subsequent purification of proteins could result in the loss of interaction if the calcium environment changes as it is not currently known if CBLs and CIPKs still interact if basal calcium is lost and this may vary depending on the interaction in question. For example, AtCBL10 requires the salt induce calcium signal to relocalise and interact with AtCIPK24 and it is not known if this interaction is abolished post calcium signal. Due to the various potential ways that interactions can be interfered, as outlined above, it is integral to use multiple interaction tests for CBLs and CIPKs to have any confidence in the demonstrated interactions.

Bimolecular fluorescence complementation is an effective method to interrogate interactions between CBLs and CIPKs from *M. polymorpha* however there are other options that have recently become available in *M. polymorpha* that may have to be considered. Despite the fact that CBL-A localises to the tonoplast membrane in *M. polymorpha* (Fig.3.2) such localisation was not detectable in *N. benthamiana* when testing interactions (Fig.5.3; Fig.5.4; Fig.5.5). Similarly the localisation sequence of CBL-A was not detectable in the primary protein sequence of CBL-A despite the obvious phylogenetic grouping into the tonoplast localising group. This should not affect interactions as generally CIPKs are recruited by CBLs to their subcellular localisation but does show distinct difference between heterologous systems which are typically used to investigate protein interactions. Of course this may be a unique issue for *M. polymorpha* due evolutionary distance. However, recently transient transformation protocols are being reported that could be used to circumvent this issue and make *M. polymorpha* the model system to test for interactions between proteins from early diverging land plants using bimolecular fluorescence complementation (Iwakawa et al. 2021). While BiFC in *N. benthamiana* is a good option for testing interactions between proteins for other angiosperms, BiFC carried out in *M. polymorpha* could become the best system for testing interactions between proteins

of interest in early diverging plants such as *P. patens* and *M. polymorpha* as it is likely more representative of interactions that could take place. For the CBL-CIPK system it may even be possible to test interactions with BiFC in *M. polymorpha* and also place the plants under under conditions that trigger calcium signals such as salt stress and see if different interactions can take place. While this opens a lot of options this may also add a lot of complexity for testing interaction and therefore it may simply be easier to knockout genes of CIPKs and CBLs in *M. polymorpha* and phenotype extensively to detect overlap of phenotypes between CBL and CIPK knockout to determine pathways and functions at the same time.

7.5 *Marchantia polymorpha* is a good system to decode the CDPK, CBL and CIPK pathways

Phylogeny has been previously used as a tool to indicate functional redundancy and function of CDPKs in *A. thaliana*. Some studies have successfully managed to use phylogeny to indicate functionally redundant genes so that they can knock-out genes that express proteins of overlapping function to determine pathways (Boudsocq et al. 2010). Typically functional redundancy can be determined by very close relation in phylogeny but the phylogeny indicated only single members of CDPKs in subgroups II, III and IV (Fig.3.1). Albeit subgroup III CDPKs -D/F can only be considered single members due to very high conserved identity and as they are encoded on the sex chromosomes of *M. polymorpha* and are therefore not considered to both be present in the same adult plant. If functions are found in the diploid stage of the lifecycle then these proteins will be considered two seperate genes and protein products which would also imply functional redundancy (Fig.3.1). The 3 CDPKs of subgroup I in *M. polymorpha* (-A/B/G) may have some functional redundancy but may also imply that the key function of calcium decoding components in early diverging land plants is drought and salt responses, as this is the main function of subgroup I CDPKs as defined in *A. thaliana*. Therefore the phylogeny results of *M.*

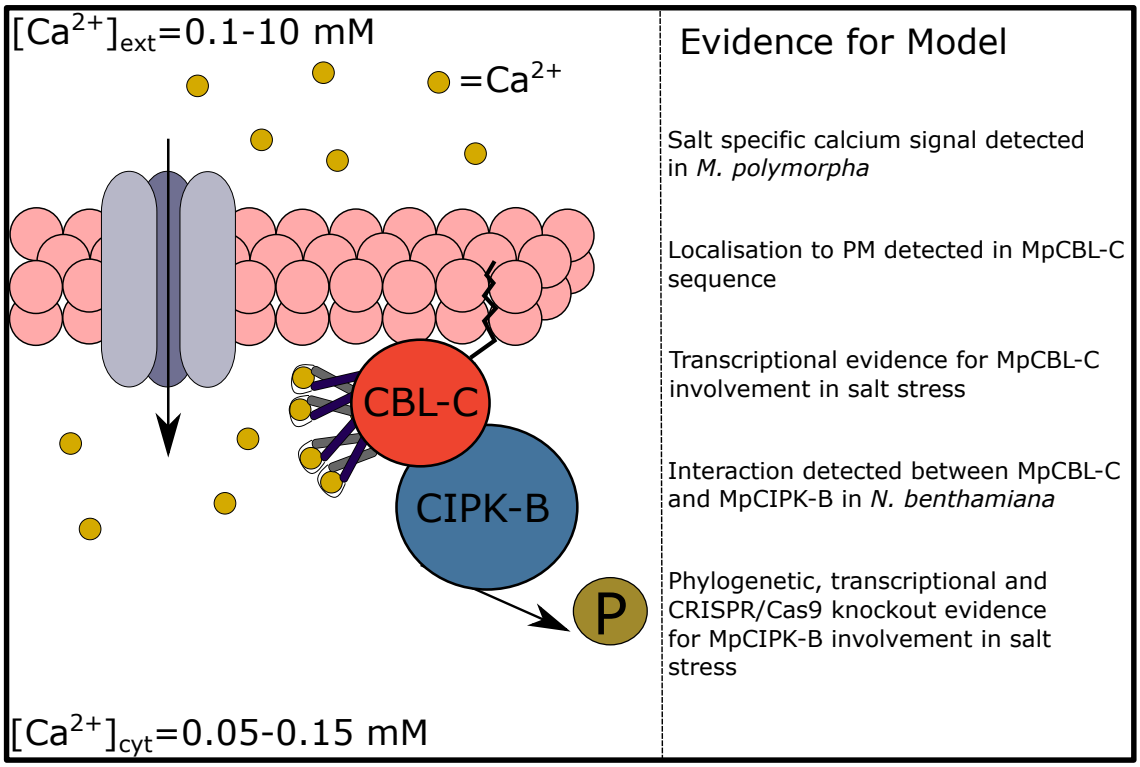


Figure 7.3: Proposed model of a salt responsive calcium decoding pathway from *M. polymorpha*. Based on the evidence proposed in this thesis there may be a salt induced calcium signal triggering a MpCBL which activates MpCIPK-B to phosphorylate a downstream protein to assist with salt tolerance in *M. polymorpha*. This is based on evidence of a salt induced calcium signal (Fig 4.1), and transcriptional data implying that MpCBL-C and MpCIPK-B are salt responsive (Section 4.2.6, 4.2.7, 4.2.8, 4.2.9), and that they interact in *N. benthamiana* (Section 5.2.2). Phylogeny implies that MpCIPK-B may be a SOS2 homologue (Section 3.2.3) and knockout mutants of *cipk-b* demonstrate reduced tolerance to salt. It is possible that MpCBL-C may be a SOS3 homologue as it has the same localisation although it does not phylogenetically group with AtCBL4/SOS3 (Section 3.2.2).

polymorpha CDPKs implies low functional redundancy in this system compared to other plants. Also the fact that subgroup II and subgroup III CDPKs have single representatives when these groups have expanded to be considered to have futher subdivided into 'a' and 'b' subsubgroups in angiosperms may offer the chance to look at the early evolutionary function of this group. Analysis of these proteins and proteins in higher plants could provide for further understanding how subgroups diverge in function to develop our evolutionary understanding of neofunctionalisation of CDPKs. *P. patens* has 21 CDPKs; 5 in subgroup I, 6 in subgroup II, 6 in subgroup III and 4 in subgroup IV. All of the PpCDPKs also form outgroups to the

AtCPKs and therefore also do not fall into 'a' and 'b' subgroups despite being higher in number than those present in *M. polymorpha* (Hamel et al. 2014). Therefore comparison between *M. polymorpha* and *P. patens* CDPKs offers a good opportunity to investigate the expansion of ancient CDPKs likely due to whole genome duplication. Research into this area would require *M. polymorpha* as a starting point due to its low functional redundancy and demonstrates good opportunities for study in this plant.

Phylogeny has been a useful tool in indicating function of CIPKs in plants and this also holds true for *M. polymorpha*. CIPK phylogeny indicated that CIPK-B was most closely related to AtCIPK24 and AtCIPK8 which are involved in drought and salt stress responses including the Salt Overly Sensitive pathway (Fig.3.1, Fig.7.3). The *cipk-b* knockout has confirmed this functional role although we cannot yet determine if CIPK-B operates as a AtCIPK24 homologue in salt response or simply acts in a similar functional role to either salt or drought. However CIPK-A also seemed related to this group albeit as an outgroup and therefore it could also be involved in drought or salt responses if we rely on phylogeny (Fig.3.1, Fig.7.3). The *cipk-b* did have increased sensitivity to salt (Fig.6.5) and which matched the phenotype of *sos2-1* but it cannot yet be determined if it is functionally redundant until the *cipk-a* lines are characterised. Previously it was thought that only AtCIPK24 was the salt responsive CIPK with AtCBL4/10 however AtCIPK8 was also characterised as part of Salt Overly Sensitive pathway through interaction with AtCBL10 to control SOS1 activity (Yin et al. 2019). However the characterisation of AtCIPK8 means that in salt pathways there may be functional redundancy that doesnt occlude salt sensitivity phenotypes. The phylogeny could also be interpreted as CIPK-A showing similarity to the outgroup of AtCIPK8/24 including AtCIPK26/3/23/9. AtCIPK26 is part of the ABA network through its regulation of 'Keep On Going' (Lyzenga et al. 2017) and also is involved in the control of reactive oxygen species generation through AtRbohF (Drerup et al. 2013). AtCIPK3 is involved in ABA responses involved in seed germination ABR1 (Sanyal et al. 2017a), an APETALA2 domain

transcription factor, and CIPK-A could also have functions here as ABA is already known to be linked to gemmae dormancy. AtCIPK23 is involved in a number of root nutrient transporters involved in balance of ions of K^+ (Liu et al. 2013), NH_4^+ and blue light dependent stomatal opening which also has implications in drought responses (Sánchez-Barrena et al. 2020; Inoue et al. 2020). AtCIPK9 is also involved in potassium responses though AtHAK5 and KAT2 (Ragel et al. 2015; Ronzier et al. 2014). Therefore mutants of CIPK-A would be expected to be involved in ionic balancing responses, either through potassium paticularly in low potassium conditions, or ABA responsive pathways involved in gemmae dormancy. Many of these responses are still involved in drought and salt dependent reponses such as ionic balancing and ABA as both pathways play key roles in drought tolerance through various mechanisms.

The EF hand analysis presented in this study may offer unprecedented resolution to phylogenetic analysis by focussing analysis on functional domains. The EF hand analysis when applied to the CDPKs from *M. polymorpha* demonstrated that each of the EF hands in the MpCDPKs was most similar to EF hands from AtCPKs from the same subgroup, which confirmed the subgrouping from the phylogeny using the MpCDPK full sequences. Therefore EF hands of CDPKs seems to differ significantly between subgroup which may also provide implications that EF hand selection and evolution is key to neofunctionalisation and development of new subgroup and function (Chapter 3.2.4). Many more studies would be required to determine if this is true but this could be carried out by looking at the development of subsubgroups 'a' and 'b' in subgroup II and III of the CDPKs. While it is no surprise that the most similar EF hands to the first EF hands in CBLs were other CBLs due to the non-canonical nature of the 14 aa loop, all other MpCBL EF hands were also most similar to AtCBL EF hands which implies early divergence of EF hands between different protein families. The EF hand analysis from the MpCDPKs meant that a 'most identical' AtCDPK could be determined which may imply that the MpCDPK is involved in the same function as the AtCDPK identified as most

similar. MpCBLs could not identify unique AtCBLs that are 'most identical' to each MpCBL and therefore the 10 AtCBLs did not provide the required resolution to this form of study (Chapter 3.2.4). However phylogeny is not typically used to provide potential function for CBLs beyond determining likely localisation. Two of the CBLs in this study were found to be very close in terms of phylogeny (CBL-B/C) and yet the expression data under increasing salt demonstrated that *CBL-C* is down regulated in a dosage dependent manner but *CBL-B* was not (Fig.4.8), so it is still unlikely that they are functionally redundant.

Based on all of the data provided in this thesis there is little indication of extensive functional redundancy in calcium decoding proteins of CDPK, CBL and CIPK families in *M. polymorpha*. CDPKs may have the highest evidence for potential functional redundancy but only in subgroup I. Therefore *M. polymorpha* offers previously unprescedented potential to understand the function of its calcium decoding toolkit as it may be possible to carry out genetic knock-out of all of the calcium decoding genes and ascribe function. This model can therefore be used to deeply interrogate the mechanistic action of these different calcium decoding protein families. *M. polymorpha* also provides opportunity to develop evolutionary insight into how these different protein families have expanded and neofunctionalised across the green lineage through macroevolutionary comparisons with studied plants such as angiosperms. Similarly comparisons to other bryophytes with much wider expansion of the calcium decoding network such as *P. patens* may provide insights to key steps in the evolutionary process such as genome duplication.

Appendix A

FamilyFinder

```
FamilyFinder [input] [protein family] [database]
```

Listing A.1: executable Family Finder

```
#!/bin/bash
# program to pblast an input against protein databases
# three arguments required
# (1) path/to/inputfile.extension [fasta format]
# (2) gene family name with no spaces
# (3) path/to/database

#add modules for program use
module add ncbi-blast/2.9.0+/gcc
module add python/anaconda/2019.3/3.7
module add mafft/7.310
module add trimal/1.2

#add input through arguments
genome="$3"
input="$1"
comp=~/script/comparison/AtCAM_protein.txt
fam="$2"

#blastp protein sequences against predicted proteins, output as
   sequence ID and sequence
```

```
21 blastp -db $genome -query $input -evalue 1e-50 -outfmt "6 sseqid
       sseq" |
22
23 #add > to each ID
24 sed 's/^/>/' |
25
26 #sort by unique ID
27 sort -u |
28
29 #remove gaps from alignment
30 sed -e "s/-//g" > ~/script/output/pipe.txt
```

Listing A.2: executable Family Finder

```
1  #!/usr/bin/env python
2  import os
3  import sys
4  pipe = open("/gpfs/home/ebh14wwu/script/output/pipe.txt", "r")
5  length = open(sys.argv[1], "r")
6  seqs = set()
7
8  min_seq_length = 51
9  max_seq_length= 666
10
11 for line in length:
12     if 50 < len(line) < min_seq_length:
13         min_seq_length = len(line)
14
15 for line in length:
16     if max_seq_length < len(line):
17         max_seq_length = len(line)
18
19 for line in pipe:
20     if max_seq_length + 20 > len(line) > min_seq_length - 20:
21         seqs.add(line)
22
```

```
with open("/gpfs/home/ebh14wwu/script/output/CIPKoutput.txt", "w")
    as output:
     for seq in seqs:
         output.write(seq.replace("\t", "\n"))

output.close()
```

Listing A.3: executable Family Finder

```

#multiple sequence alignment
mafft --globalpair --retree 100 --maxiterate 1000  ~/script/output/
    CIPKoutput.txt > ~/script/output/$2align.txt

#trimal1.2 to trim for realignment
trimal -in ~/script/output/$2align.txt -out ~/script/output/$2trim.
    txt -automated1

#Modified from Alon Shaiber for isolating unique sequences , github
    .io 2019
sed -e "s/-//g" ~/script/output/$2trim.txt |\
sed '/^>/s/$/@/' |\
sed 's/^>/#/' |\
tr -d '\n' | tr "#" "\n" | tr "@" "\t" |\
sort -u -f |\
sed -e 's/^/>/' |\
tr '\t' '\n/' |
tail -n +2 > ~/script/output/$2trim2.txt

#merge inputs
cat ~/script/output/$2trim2.txt $input $comp > ~/script/output/$2.
    txt

#realign
mafft --globalpair --retree 100 --maxiterate 1000 --treeout ~/
```

```
script/output/$2.txt > ~/script/output/$2_align.txt
```

Listing A.4: executable Family Finder

Bibliography

Aki, S. S., T. Mikami, S. Naramoto, R. Nishihama, et al. (2019). "Cytokinin Signaling Is Essential for Organ Formation in Marchantia polymorpha". In: *Plant and Cell Physiology* 60, pp. 1824–1854. DOI: 10.1093/pcp/pcz100.

Albert, N. W., A. H. Thrimawithana, T. K. McGhie, W. A. Clayton, et al. (2018). "Genetic analysis of the liverwort Marchantia polymorpha reveals that R2R3MYB activation of flavonoid production in response to abiotic stress is an ancient character in land plants". In: *New Phytologist*. ISSN: 1469-8137.

Albrecht, V., O. Ritz, S. Linder, K. Harter, et al. (2001). "The NAF domain defines a novel protein-protein interaction module conserved in Ca2+-regulated kinases". In: *EMBO Journal* 20, pp. 1051–1063.

Albrecht, V., S. Weinl, D. Blazevic, C. D'Angelo, et al. (2003). "The calcium sensor CBL1 integrates plant responses to abiotic stresses". In: *The Plant Journal*. DOI: 10.1046/j.1365-313X.2003.01892.x.

Ali, R., R. E. Zielinski, and G. Berkowitz (2006). "Expression of plant cyclic nucleotide-gated cation channels in yeast". In: *Journal of Experimental Biology* 57, pp. 125–138.

Allan, C., R. Morris, and C.-N. Meisrimler (2022). "Encoding, transmission, decoding, and specificity of calcium signals in plants". In: *Journal of Experimental Botany* 73, pp. 3372–3385.

Allen, G. J., S. P. Chu, K. Schumacher, C. T. Shimazaki, et al. (2000). "Alteration of stimulus-specific guard cell calcium oscillations and stomatal closing in Arabidopsis det3 mutant." In: *Science* 289.5488, pp. 2338–2342.

Allen, G. J., S. P. Chu, C. L. Harrington, K. Schumacher, et al. (2001). "A defined range of guard cell calcium oscillation parameters encodes stomatal movements." In: *Nature* 411.6941, pp. 1053–1057.

Althoff, F., S. Kopischke, O. Zobell, K. Ide, et al. (2014). "Comparison of the MpEF1α and CaMV35 promoters for application in Marchantia polymorpha overexpression studies". In: *Transgenic Research* 23, pp. 235–244. DOI: 10.1007/s11248-013-9746-z.

Ares, A., M. Itouga, Y. Kato, and H. Sakakibara (2018). "Differential Metal Tolerance and Accumulation Patterns of Cd, Cu, Pb and Zn in the Liverwort Marchantia polymorpha L." In: *Bull Environ Contam Toxicol* 100, pp. 444–450. DOI: 10.1007/s00128-017-2241-0.

Asano, T., N. Hayashi, M. Kobayashi, N. Aoki, et al. (2012). "A rice calcium-dependent protein kinase OsCPK12 oppositely modulates salt-stress tolerance and blast disease resistance". In: *Plant Journal* 69.26-36. DOI: 10.1111/j.1365-313X.2011.04766.x.

Barnes, C. R. and W. J. G. Land (1908). "Bryology Papers II. The origin of the cupule of Marchantia". In: *Botany Gazette* 46, pp. 401–409.

Batistič, O. and J. Kudla (2004). "Integration and channeling of calcium signaling through the CBL calcium sensor/CIPK protein kinase network". In: *Planta*. DOI: 10.1007/s00425-004-1333-3.

Batistič, O., M. Rehers, A. Akermann, K. Schlucking, et al. (2012). "S-acylation-dependent association of the calcium sensor CBL2 with the vacuolar membrane is essential for proper abscisic acid responses". In: *Cell Research* 22, pp. 1155–1168. DOI: doi:10.1038/cr.2012.71.

Batistič, O., N. Sorek, S. Schultke, S. Yalovsky, et al. (2008). "Dual Fatty Acyl Modification Determines the Localization and Plasma Membrane Targeting of CBL/CIPK Ca2+ Signaling Complexes in Arabidopsis." In: *The Plant Cell Online* 20.5, pp. 1346–1362. DOI: doi:10.1105/tpc.108.058123.

Batistič, O., R. Waadt, L. Steinhorst, K. Held, et al. (2010). "CBL-mediated targeting of CIPKs facilitates the decoding of calcium signals emanating from distinct cellular stores". In: *The Plant Journal* 61.2, pp. 211–222. DOI: https://doi.org/10.1111/j.1365-313X.2009.04045.x.

Beckmann, L., K. H. Edel, O. Batistič, and J. Kudla (2016). "A calcium sensor - protein kinase signaling module diversified in plants and is retained in all lineages of Bikonta species". In: *Scientific Reports* 6.31645. DOI: 10.1038/srep31645.

Bender, K. W., R. K. Blackburn, J. Monaghan, P. Derbyshire, et al. (2017). "Autophosphorylation-based calcium (Ca2+) sensitivity priming and Ca2+/Calmodulin inhibition of Arabidopsis thaliana Ca2+-dependent protein kinase 28 (CPK28)". In: *The Journal of Biological Chemistry* 292.10, pp. 1988–4002. DOI: 10.1074/jbc.M116.763243.

Berrie, G. K. (1960). "The chromosome numbers of liverworts (Hepaticae and Anthocerotae)". In: *Trans. Brit. Bryol. Soc.* 3, pp. 688–705.

Beyhl, D., S. Hortensteiner, E. Martinoia, E. E. Farmer, et al. (2009). "The fou2 mutation in the major vacuolar cation channel TPC1 confers tolerance to inhibitory luminal calcium". In: *Plant Journal* 58, pp. 715–723.

Birchler-Causse, H. (1993). *Marchantia L. The European and African Taxa,* vol. 45. Berlin, Stuttgart: J. Cramer.

Blumwald E. Aharon, G. S. and M. P. Apse (2000). "Sodium Transport in Plant Cells". In: *BBA Biomembranes* 1465.140-151. DOI: 10.1016/S0005-2736(00)00135-8.

Boudsocq, M., M. J. Droillard, L. Regad, and C. Lauriere (2012). "Characterization of Arabidopsis calcium-dependent protein kinases: activated or not by calcium?" In: *Biochemical Journal* 447.2, pp. 291–299. DOI: 10.1042/BJ20112072.

Boudsocq, M. and J. Sheen (2013). "CDPKs in immune and stress signaling." In: *Trends in Plant Science* 18.1, pp. 30–40. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3534830/pdf/nihms402326.pdf.

Boudsocq, M., M. R. Willmann, M. McCormack, H. Lee, et al. (2010). "Differential innate immune signalling via Ca(2+) sensor protein kinases". In: *Nature* 464.7287, pp. 418–22. ISSN: 0028-0836. DOI: 10.1038/nature08794. URL: http://www.nature.com/articles/nature08794.pdf.

Bowler, C. and R. Fluhr (2000). "The role of calcium and activated oxygens as signals for controlling cross-tolerance." In: *Trends in Plant Science* 5.6, pp. 241–246.

Bowman, J. L. (2016). "A Brief History of Marchantia from Greece to Genomics." In: *Plant & Cell Physiology* 57, pp. 210–229.

Bowman, J. L., T. Araki, M. A. Arteaga-Vazquez, F. Berger, et al. (2016). "The Naming of Names: Guidlines for Gene Nomenclature in Marchantia". In: *Plant Cell Physiology* 57, pp. 257–261.

Bowman, J. L., S. K. Floyd, and K. Sakakibara (2007). "Green genes-comparative genomics of the green branch of life". In: *Cell* 129.2, pp. 229–234.

Bowman, J. L., T. Kohchi, K. T. Yamato, J. Jenkins, et al. (2017). "Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome." In: *Cell* 171.2, pp. 287–304.

Brailoiu, E., D. Churamani, X. Cai, M. G. Schrlau, et al. (2009). "Essential requirement for two-pore channel 1 in NAADP-mediated calcium signalling". In: *Journal of Cell Biology* 186, pp. 201–209.

Brandt, B., S. Munemasa, C. Wang, D. Nguyen, et al. (2015). "Calcium specificity signaling mechanisms in abscisic acid signal transduction in Arabidopsis guard cells." In: *eLIFE* 4, e03599. DOI: 10.7554/eLife.03599.

Calcraft, P. J., M. Ruas, Z. Pan, X. Cheng, et al. (2009). "NAADP mobilises calcium from acidic organelles through two-pore channels". In: *Nature* 459, pp. 596–600.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, et al. (2009). "BLAST+: architecture and applications". In: *BMC Bioinformatics* 10.421. DOI: 10.1186/1471-2105-10-421.

Cang, C., Y. Zhou, B. Nevarro, Y. J. Seo, et al. (2013). "mTOR regulates lysosomal ATP-sensitive two-pore Na+ channels to adapt to metabolic state". In: *Cell* 152, pp. 778–790.

Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón (2009). "trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses". In: *Bioinformatics* 25.15, pp. 1972–1973. DOI: 10.1093/bioinformatics/btp348.

Catalá, R., E. Santos, J. M. Alonso, J. R. Ecker, et al. (Dec. 2003). "Mutations in the Ca2+/H+ Transporter CAX1 Increase CBF/DREB1 Expression and the Cold-Acclimation Response in Arabidopsis". In: *The Plant Cell* 15.12, pp. 2940–2951. ISSN: 1040-4651. DOI: 10.1105/tpc.015248. eprint: https://academic.oup.com/plcell/article-pdf/15/12/2940/35429538/plcell_v15_12_2940.pdf. URL: https://doi.org/10.1105/tpc.015248.

Chandran, V., E. J. Stollar, K. Lindorff-Larsen, J. F. Harper, et al. (2006). "Structure of the Regulatory Apparatus of a Calcium-dependent Protein Kinase (CDPK): A Novel Mode of Calmodulin-target Recognition". In: *Journal of Molecular Biology* 357.2, pp. 400–410.

Charpentier, M., J. Sun, V. Martins, G. V. Radhakrishnan, et al. (2016). "Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations." In: *Science* 352.6289, pp. 1102–1105. DOI: 10.1126/science.aae0109..

Chen, D. H., H. P. Liu, and C. L. Li (2018). "Calcium-dependent protein kinase CPK9 negatively functions in stomatal abscisic acid signaling by regulating ion channel activity in Arabidopsis". In: *Plant Molecular Biology* 99.113-122. DOI: 10.1007/s11103-018-0805-y.

Chen, F., L. Zhang, and Z. Cheng (2017). "The calmodulin fused kinase novel gene family is the major system in plants converting Ca2+ signals to protein phosphorylation responses". In: *Scientific Reports* 7. DOI: doi:10.1038/s41598-017-03367-8.

Chen, J. and V. Mueller (2018). "Coastal climate change, soil salinity and human migration in Bangladesh". In: *Nature Climate Change* 8, pp. 981–985. DOI: 10.1038/s41558-018-0313-8.

Cheng, N. H., J. K. Pittman, J. K. Zhu, and K. D. Hirschi (2004). "The Protein Kinase SOS2 Activates the Arabidopsis H+/Ca2+ Antiporter CAX1 to Integrate Calcium Transport and Salt Tolerance*". In: *Journal of Biological Chemistry* 279.4, pp. 2922–2926. ISSN: 0021-9258. DOI: https://doi.org/10.1074/jbc.M309084200. URL: https://www.sciencedirect.com/science/article/pii/S0021925818526636.

Cheong, Y. H., K. N. Kim, G. K. Pandey, R. Gupta, et al. (2003). "CBL1, a calcium sensor that differentially regulates salt, drought, and cold responses in Arabidopsis." In: *The Plant Cell*. DOI: 10.1105/tpc.012393..

Cheong, Y. H., G. K. Pandey, J. J. Grant, O. Batistič, et al. (2007). "Two calcineurin B-like calcium sensors, interacting with protein kinase CIPK23, regulate leaf transpiration and root potassium uptake in Arabidopsis." In: *Plant Journal* 52.2, pp. 223–239.

Chiyoda, S., K. Ishizaki, H. Kataoka, K. T. Yamato, et al. (2008). "Direct transformation of the liverwort Marchantia polymorpha L. by particle bombardment using immature thalli developing from spores." In: *Plant Cell Reports* 27.9, pp. 1467–1473.

Cho, J. H., J. H. Lee, Y. K. Park, M. N. Choi, et al. (2016). "Calcineurin B-like Protein CBL10 Directly Interacts with TOC34 (Translocon of the Outer Membrane of the Chloroplasts) and Decreases Its GTPase Activity in Arabidopsis". In: *Frontiers in Plant Science*. DOI: 10.1104/pp.108.130419.

Choi, W. G., M. Toyota, S. H. Kim, R. Hilleary, et al. (2014). "Salt stress-induced Ca2+ waves are associated with rapid, long-distance root-to-shoot signaling in plants". In: *PNAS* 111, pp. 6497–6502. DOI: 10.1073/pnas.1319955111.

Christodoulou, J., A. Malmendal, J. F. Harper, and W. J. Chazin (2004). "Evidence for Differing Roles for Each Lobe of the Calmodulin-like Domain in a

Calcium-dependent Protein Kinase". In: *The Journal of Biological Chemistry* 279, pp. 29092–29100. DOI: 10.1074/jbc.M401297200. URL: http://www.jbc.org/content/279/28/29092.full.pdf.

Clapham, D. E. (2007). "Calcium Signalling". In: *Cell.* 131.6, pp. 1047–1058.

Crandall-Stotler, B. (1981). "Morphology/anatomy of hepatics and antho-cerotes". In: *Advances in Bryology*, pp. 315–398.

Crizel, R. L., E. C. Perin, I. L. Vighi, R. Woloski, et al. (2020). "Genome-wide identification, and characterization of the CDPK gene family reveal their involvement in abiotic stress response in Fragaria x ananassa". In: *Scientific Reports* 10, p. 11040. DOI: 10.1038/s41598-020-67957-9.

Day, I. S., V. S. Reddy, G. S. Ali, and A. S. N. Reddy (2002). "Analysis of EF-hand-containing proteins in Arabidopsis". In: *Genome Biology* 3.10, research0056.1–0056.24. DOI: 10.1186/gb-2002-3-10-research0056.

Dazy, M., J. F. Masfaraud, and J. F. Ferard (2009). "Induction of oxidative stress biomarkers associated with heavy metal stress in Fontinalis antipyretica Hedw". In: *Chemosphere* 75, pp. 297–302. DOI: 10.1016/j.chemosphere.2008.12.045.

De Castro, E., C. J. A. Sigrist, V. Gattiker A. Bulliard, P. S. Langendijk-Genevaux, et al. (2006). "ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins". In: *Nucl Acids Res* 1.34, pp. 362–365. DOI: https://doi.org/10.1093/nar/gkl124.

De Vriese, K., A. Costa, T. Beeckman, and S. Vanneste (2018). "Pharmacological Strategies for Manipulating Plant Ca2+ Signalling". In: *Int J Mol Sci* 19, p. 1506.

DeFalco, T. A., C. B. Marshall, K. Munro, H. Kang, et al. (2016). "Multiple Calmodulin-Binding Sites Positively and Negatively Regulate Arabidopsis CYCLIC NUCLEOTIDE-GATED CHANNEL12". In: *Plant Cell* 28, pp. 1738–1751. DOI: https://doi.org/10.1105/tpc.15.00870.

Demidchik, V., R. J. Davenport, and M. Tester (2002). "Nonselective Cation Channels in Plants". In: *Annual Review of Plant Biology* 53.67-107. DOI: 10.1146/annurev.arplant.53.091901.161540.

Dolmetsch, R. E., R. S. Lewis, C. C. Goodnow, and J. I. Healy (1997). "Differential activation of transcription factors induced by Ca2+ response amplitude and duration." In: *Nature* 386.6627, pp. 855–858.

Dolmetsch, R. E., K. Xu, and R. S. Lewis (1998). "Calcium oscillations increase the efficiency and specificity of gene expression." In: *Nature* 30.6679, pp. 933–936.

Drabikowski, W., H. Brzeska, and S. Y. Venyaminov (1984). "Tryptic fragments of calmodulin. Ca2+- and Mg2+-induced conformational changes." In: *Journal of Biological Chemistry* 257, pp. 11584–11590.

Drerup, M. M., K. Schlücking, K. Hashimoto, P. Manishankar, et al. (2013). "The Calcineurin B-Like Calcium Sensors CBL1 and CBL9 Together with Their Interacting Protein Kinase CIPK26 Regulate the Arabidopsis NADPH Oxidase RBOHF". In: *Molecular Plant* 6.2, pp. 559–569. ISSN: 1674-2052. DOI: https://doi.org/10.1093/mp/sst009. URL: http://www.sciencedirect.com/science/article/pii/S1674205214601118.

Dubeaux, G., N. Julie, E. Zelazny, and G. Vert (2018). "Metal Sensing by the IRT1 Transporter-Receptor Orchestrates Its Own Degradation and Plant Metal Nutrition". In: *Molecular Cell* 69, pp. 953–964. DOI: 10.1016/j.molcel.2018.02.009.

Edel, K. H. and J. Kudla (2015). "Increasing complexity and versatility: how the calcium signaling toolkit was shaped during plant land colonization." In: *Cell Calcium* 57.3, pp. 231–246. DOI: doi:10.1016/j.ceca.2014.10.013.

Einstein, A. (1905). "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In: *Annalen der Physik* 322.8. DOI: https://doi.org/10.1002/andp.19053220806.

Eklund, D. M., E. Kanei M. Flores-Sandoval, K. Ishizaki, R. Nishihama, et al. (2018). "An Evolutionarily Conserved Abscisic Acid Signaling Pathway Regulates Dormancy in the Liverwort Marchantia polymorpha". In: *Current Biology* 28.3691-3699. DOI: 10.1016/j.cub.2018.10.018.

Fang, H., L. Huangfu, R. Chen, P. Li, et al. (2017). "Ancestor of land plants acquired the DNA-3-methyladenine glycosylase (MAG) gene from bacteria through horizontal gene transfer." In: *Scientific Reports* 7, p. 9324. DOI: 10.1038/s41598-017-05066-w.

Feng, Z., B. Zhang, W. Ding, X. Liu, et al. (2013). "Efficient genome editing in plants using a CRISPR/Cas system". In: *Cell Research* 23, pp. 1229–1232. DOI: 10.1038/cr.2013..

Flores-Sandoval, E., D. M. Eklund, S.-F. Hong, J. P. Alvarez, et al. (2018). "Class C ARFs evolved before the origin of land plants and antagonize differentiation and developmental transitions in Marchantia polymorpha". In: *New Phytologist* 218, pp. 1612–1630. DOI: 10.1111/nph.15090.

Flores-Sandoval, E., M. D. Eklund, and J. L. Bowman (2015). "A Simple Auxin Transcriptional Response System Regulates Multiple Morphogenetic Processes in the Liverwort Marchantia polymorpha". In: *PLoS Genetics* 12, e1005900. DOI: 10.1371/journal.pgen.1005207.

Flowers, T. J. and A. R. Yeo (1986). "Ion Relations of Plants Under Drought and Salinity." In: *Australian Journal of Plant Physiology* 13.75-91. DOI: doi.org/10.1071/PP9860075.

Fraile-Escanciano, A., Y. Kamisugi, A. C. Cuming, A. Rodríguez-Navarro, et al. (2010). "The SOS1 transporter of Physcomitrella patens mediates sodium efflux in planta". In: *New Phytologist* 188, pp. 750–761. DOI: 10.1111/j.1469-8137.2010.03405.x.

Fu, L., X. Yu, and C. An (2013). "Overexpression of constitutively active OsCPK10 increases Arabidopsis resistance against Pseudomonas syringae pv. tomato and rice resistance against Magnaporthe grisea." In: *Plant Physiological Biochemistry* 73, pp. 202–210. DOI: 10.1016/j.plaphy.2013.10.004.

Furumoto, T., N. Ogawa, S. Hata, and K. Izui (1996). "Plant calcium-dependent protein kinase-related kinases (CRKs) do not require calcium for their activities". In: *FEBS Letters* 396.2-3, pp. 147–151.

Gao, Q. F., C. F. Fei, J. Y. Dong, L. L. Gu, et al. (2014). "Arabidopsis CNGC 18 is a Ca2+-permeable channel". In: *Molecular Plant* 7, pp. 739–743.

Gentry, H. R., A. U. Singer, L. Betts, C. Yang, et al. (2005). "Structural and biochemical characterization of CIB1 delineates a new family of EF-hand-containing proteins." In: *Journal of Biological Chemistry* 280.8407-15. DOI: 10.1074/jbc.M411515200.

Gong, Q., P. Li, S. Ma, S. I. Rupassara, et al. (2005). "Salinity stress adaptation competence in the extremophile Thellungiella halophila in comparison with its relative Arabidopsis thaliana". In: *The Plant Journal* 44.826-839. DOI: 10.1111/j.1365-313X.2005.02587.x.

Grabarek, Z. (2005). "Structure of a Trapped Intermediate of Calmodulin: Calcium Regulation of EF-hand Proteins from a New Perspective." In: *Journal of Molecular Biology* 346, pp. 1351–1366. DOI: 10.1016/j.jmb.2005.01.004.

Greaves, J., G. Prescott, O. Gorleku, and L. Chamberlain (2009). "The fat controller: roles of palmitoylation in intracellular protein trafficking and targeting to membrane microdomains." In: *Molecular Membrane Biology*. DOI: 10.1080/09687680802620351.

Halfter, U., M. Ishitani, and J. K. Zhu (2000). "The Arabidopsis SOS2 protein kinase physically interacts with and is activated by the calcium-binding protein SOS3." In: *Proc Natl Acad Sci U S A*. DOI: 10.1073/pnas.040577697.

Hamel, L. P., J. Sheen, and A. Seguin (2014). "Ancient signals: comparative genomics of green plant CDPKs". In: *Trends in Plant Science* 19, pp. 79–89. DOI: 10.1016/j.tplants.2013.10.009.

Harmon, A., B. Yoo, and C. McCaffery (1994). "Pseudosubstrate Inhibition of CDPK, a Protein Kinase with a Calmodulin-like Domain". In: *Biochemistry* 33, pp. 7278–7287.

Harper, J. F. (2001). "Dissecting calcium oscillators in plant cells". In: *Trends in Plant Science* 6.9, pp. 395–397. DOI: https://doi.org/10.1016/S1360-1385(01)02023-4.

Harper, J. F., J. F. Huang, and S. J. Lloyd (1994). "Genetic Identification of an Autoinhibitor in CDPK, a Protein Kinase with a Calmodulin-like domain". In: *Biochemistry* 33, pp. 7267–7277.

Harris, B. J., C. J. Harrison, A. M. Hetherington, and T. A. Williams (2020). "Phylogenomic Evidence for the Monophyly of Bryophytes and the Reductive Evolution of Stomata". In: *Current Biology* 30, pp. 2001–2012. DOI: 10.1016/j.cub.2020.03.048.

Hashimoto, K., C. Eckert, U. Anschuetz, M. Scholz, et al. (2012). "Phosphorylation of Calcineurin B-like (CBL) calcium sensor proteins by their CBL-interacting protein kinases (CIPKs) is required for full activity of CBL–CIPK complexes towards their target proteins". In: *Journal of Biological Chemistry* 287, pp. 7956–7968.

Hassan, M. H., Y. Zhang, G. Yuan, G. A. Tuskan, et al. (2021). "Construct design for CRISPR/Cas-based genome editing in plants". In: *Trends in Plant Science* 26, pp. 1133–1152. DOI: 10.1016/j.tplants.2021.06.015.

Haswell, E. S. and E. M. Meyerowitz (2006). "MscS-like proteins control plastid size and shape in Arabidopsis thaliana". In: *Current Biology* 16, pp. 1–11.

Haswell, E. S., R. Peyrinnet, H. Bardier-Brygoo, E. M. Meyerowitz, et al. (2008). "Two MscS Homologs provide mechanosensitive channel activities in Arabidopsis root". In: *Current Biology* 18, pp. 730–734.

Held, K., C. Pascaud F. Eckert, P. Gajdanowicz, K. Hashimoto, et al. (2011). "Calcium-dependent modulation and plasma membrane targeting of the AKT2 potassium channel by the CBL4/CIPK6 calcium sensor/protein kinase complex". In: *Cell Research* 21.1116-1130. DOI: 10.1038/cr.2011.50.

Hernández-García, J., R. Sun, A. Serrano-Mislata, K. Inoue, et al. (2021). "Coordination between growth and stress responses by DELLA in the liverwort Marchantia polymorpha." In: *Current Biology* 31, 3678–3686e. DOI: 10.1016/j.cub.2021.06.010..

Hetherington, A. and C. Brownlee (2004). "The generation of Ca(2+) signals in plants". In: *Annual Review of Plant Biology* 55, pp. 401–427. DOI: 10.1146/annurev.arplant.55.031903.141624.

Hook, S. and A. R. Means (2001). "Ca2+/CaM Dependent Kinases: From Activation to Function". In: *Annual Review of Pharmacology and Toxicology* 41, pp. 471–505. DOI: https://doi.org/10.1146/annurev.pharmtox.41.1.471.

Horecka, J. and A. M. Chu (2017). "Yeast Colony PCR: It doesn't get any easier than this! V.1". In: *Protocols.io*. DOI: dx.doi.org/10.17504/protocols.io.gzwbx7e.

Hua, B., R. W. Mercier, R. E. Zielinski, and G. A. Berkowitz (2003a). "Functional interaction of calmodulin with a plant cyclic nucleotide gated cation channel". In: *Plant Physiology and Biochemistry* 41.11-12, pp. 945–954. DOI: https://doi.org/10.1016/j.plaphy.2003.07.006.

Hua, W., S. Liang, and Y. T. Lu (2003b). "A tobacco (Nicotiana tabaccum) calmodulin-binding protein kinase, NtCBK2, is regulated differentially by calmodulin isoforms". In: *Biochemical Journal* 376.1, pp. 291–302. DOI: 10.1042/bj20030736. URL: http://www.biochemj.org/content/ppbiochemj/376/1/291.full.pdf.

Huang, F., J. Luo, T. Ning, W. Cao, et al. (2017). "Cytosolic and Nucleosolic Calcium Signaling in Response to Osmotic and Salt Stresses Are Independent of Each Other in Roots of Arabidopsis Seedlings". In: *Frontiers Plant Science* 8.1648. DOI: 10.3389/fpls.2017.01648.

Hubbard, K. E., R. S. Siegel, G. Valerio, B. Brandt, et al. (2012). "Abscisic acid and CO2 signalling via calcium sensitivity priming in guard cells, new CDPK mutant phenotypes and a method for improved resolution of stomatal stimulus-response analyses." In: *Ann Bot.* 109.1, pp. 5–17. DOI: 10.1093/aob/mcr252.

Husain, S., R. Munns, and A. G. Condon (2003). "Effect of sodium exclusion trait on chlorophyll retention and growth of durum wheat in saline soil." In: *Australian Journal of Agricultural Research* 54.589-597. DOI: doi.org/10.1071/AR03032.

Iida, K., J. Teng, T. Tada, A. Saka, et al. (2007). "Essential, completely conserved glycine residue in the domain III S2-S3 linker of voltage-gated calcium channel alpha1 subunits in yeast and mammals," in: *Journal of Biological Chemistry* 282.35, pp. 25659–25667.

Inoue, K., R. Nishihama, and T. Kohchi (2019). "Phytochrome and Light Signaling in Marchantia." In: *Methods Molecular Biology* 2026, pp. 215–223. DOI: 10.1007/978-1-4939-9612-4_18.

Inoue, S. I., E. Kaiserli, X. Zhao, T. Waksman, et al. (2020). "CIPK23 regulates blue light-dependent stomatal opening in Arabidopsis thaliana". In: *The Plant Journal* 104, pp. 679–692. DOI: doi.org/10.1104/pp.19.01084.

Irifune, K., K. Ono, M. Takahashi, H. Murakami, et al. (1996). "Stable transformation of cultured cells of the liverwortMarchantia polymorpha by particle bombardment." In: *Transgenic Research* 5, pp. 337–341. DOI: 10.1007/BF01968943.

Ishitani, M., J. Liu, U. Halfter, C. S. Kim, et al. (2000). "SOS3 Function in Plant Salt Tolerance Requires N-Myristoylation and Calcium Binding". In: *The Plant Cell* 12.9. DOI: 10.1105/tpc.12.9.1667.

Ishizaki, K., S. Choiyoda, K. T. Yamato, and T. Kohchi (2008). "Agrobacterium-mediated transformation of the haploid liverwort Marchantia polymorpha L and an emerging model for plant biology." In: *Plant Cell Physiology* 49.7, pp. 1084–1091.

Ishizaki, K., Y. Johzuka-Hisatomi, S. Ishida, S. Iida, et al. (2013a). "Homologous Recombination-Mediated Gene Targeting in the Liverwort Marchantia polymorpha L." In: *Scientific Reports* 1532. DOI: https://doi.org/10.1038/srep01532.

Ishizaki, K., M. Mizutani, M. Shimamura, A. Masuda, et al. (2013b). "Essential role of the E3 ubiquitin ligase nopperabo1 in schizogenous intercellular space formation in the liverwort Marchantia polymorpha." In: *Plant Cell* 10, pp. 4075–4084. DOI: 10.1105/tpc.113.117051..

Ito, T., M. Nakata, J. Fukuzawa, S. Ishida, et al. (2010). "Alteration of Substrate Specificity: The Variable N-Terminal Domain of Tobacco Ca2+-Dependent Pro-

tein Kinase Is Important for Substrate Recognition". In: *The Plant Cell* 22.5, pp. 1592–1604. DOI: https://doi.org/10.1105/tpc.109.073577.

Iwakawa, H., K. Melkonian, T. Schlüter, H. W. Jeon, et al. (2021). "Agrobacterium-Mediated Transient Transformation of Marchantia Liverworts". In: *Plant and Cell Physiology*, pcab126. DOI: 10.1093/pcp/pcab126.

James, R. A., A. R. Rivelli, and S. Munns R. von Caemmerer (2002). "Factors affecting CO2 assimilation, leaf injury and growth in salt-stressed durum wheat". In: *Functional Plant Biology* 29.13931403. DOI: 10.1071/FP02069.

Jamil, A., S. Riaz, M. Ashraf, and M. R. Foolad (2011). "Gene Expression Profiling of Plants under Salt Stress". In: *Critical Reviews in Plant Science* 435-458. DOI: 10.1080/07352689.2011.605739.

Ji, H., J. M. Pardo, G. Batelli, M. J. Van Oosten, et al. (2013). "The Salt Overly Sensitive (SOS) Pathway: Established and Emerging Roles". In: *Molecular Plant* 6.2, pp. 275–286. DOI: https://doi.org/10.1093/mp/sst017.

Jiang, W., H. Zhou, H. Bi, M. Fromm, et al. (2013). "Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice". In: *Nucleic acid Research* 41, e188. DOI: 10.1093/nar/gkt780.

Johnson, D. R., R. S. Bhatnagar, L. J. Knoll, and J. I. Gordon (1994). "Genetic and biochemical studies of protein N-myristoylation". In: *Annu Rev Biochem.* DOI: 10.1146/annurev.bi.63.070194.004253..

Jones, V. A. and L. Dolan (2017). "MpWIP regulates air pore complex development in the liverwort Marchantia polymorpha." In: *Development* 144, pp. 1472–1476. DOI: 10.1242/dev.144287..

Kanazawa, T., A. Era, N. Minamino, Y. Shikano, et al. (2016). "SNARE Molecules in Marchantia polymorpha : Unique and Conserved Features of the Membrane Fusion Machinery". In: *Plant and Cell Physiology* 57, pp. 307–324. DOI: 10.1093/pcp/pcv076.

Kato, H., K. Ishizaki, M. Kouno, M. Shirakawa, et al. (2015). "Auxin-Mediated Transcriptional System with a Minimal Set of Components Is Critical for Mor-

phogenesis through the Life Cycle in Marchantia polymorpha". In: *PLoS Genetics* 11, e1005365. DOI: 10.1371/journal.pgen.1005365.

Kim, B. G., R. Waadt, Y. H. Cheong, G. K. Pandey, et al. (2007). "The calcium sensor CBL10 mediates salt tolerance by regulating ion homeostasis in Arabidopsis". In: *Plant Journal* 52, pp. 473–484. DOI: https://doi.org/10.1111/j.1365-313X.2007.03249.x.

Kim, S. H., D. H. Woo, J. M. Kim, S. Y. Lee, et al. (2011). "Arabidopsis MKK4 mediates osmotic-stress response via its regulation of MPK3 activity." In: *Biochemical and Biophysical Research Communications* 412.150-154. DOI: 10.1016/j.bbrc.2011.07.064.

Kisiel, K., H. Miwa, and I. J. Odrzykoski (2011). "Taxonomic identification of chloroplast genome of Marchantia polymorpha using DNA barcode sequences." In: *In Fourth International Barcode of Life Conference*, A43.

Kleist, T., A. Spenceley, and S. Luan (2014). "Comparative phylogenomics of the CBL-CIPK calcium-decoding network in the moss Physcomitrella, Arabidopsis, and other green lineages." In: *Frontiers in Plant Science* 5.

Knight, H. and M. R. Knight (2001). "Abiotic stress signalling pathways: Specificity and cross-talk". In: *Trends in Plant Science* 6.6, pp. 262–267.

Knight, H., A. J. Trewavas, and M. R. Knight (1997). "Calcium signalling in Arabidopsis thaliana responding to drought and salinity". In: *The Plant Journal* 12.1067-1078. DOI: 10.1046/j.1365-313X.1997.12051067.x.

Knight, M. R., A. K. Campbell, S. M. Smith, and A. Trewavas (1991). "Transgenic plant aequorinreports the effects of touch and cold-shock and elicitors on cytoplasmic calcium". In: *Nature* 352, pp. 524–526.

Kolukisaoglu, U., S. Weinl, D. Blazevic, O. Batistič, et al. (2004). "Calcium Sensors and Their Interacting Protein Kinases: Genomics of the Arabidopsis and Rice CBL-CIPK Signaling Networks". In: *Plant Physiology* 134, pp. 43–58. DOI: 10.1104/pp.103.033068..

Kubota, A., K. Ishizaki, M. Hosaka, and T. Kohchi (2013). "Efficient Agrobacterium-Mediated Transformation of the Liverwort Marchantia polymorpha Using Regenerating Thalli". In: *Bioscience, Biotechnology, and Biochemistry* advpub. DOI: https://doi.org/10.1271/bbb.120700.

Kumar, S., M. Mazumder, N. Gupta, S. Chattopadhyay, et al. (2016). "Crystal structure of Arabidopsis thaliana calmodulin7 and insight into its mode of DNA binding". In: *FEBS Letters* 590.17. DOI: 10.1002/1873-3468.12349.

Kuo, M. M. C., A. Baker K, L. Wong, and S. Choe (2007). "Dynamic Oligomeric Conversions of the cytoplasmic RCK Domains Mediate MthK Potassium Channel Activity". In: *PNAS* 13.7, pp. 2151–2156. DOI: https://doi.org/10.1073/pnas.0609085104.

Laenen, B., B. Shaw, H. Schneider, B. Goffinet, et al. (2014). "Extant diversity of bryophytes emerged from successive post-Mesozoic diversification bursts." In: *Nature Communication* 5, p. 5134. DOI: 10.1038/ncomms6134.

Larebeke, N. V., G. Engler, M. Holsters, S. van der Elsacker, et al. (1974). "Large plasmid in Agrobacterium tumefaciens essential for crown gall-inducing ability." In: *Nature* 252, pp. 169–170. DOI: 10.1038/252169a0.

Lee, S. C., W. Z. Lan, B. G. Kim, L. Li, et al. (2007). "A protein phosphorylation/dephosphorylation network regulates a plant potassium channel". In: *PNAS* 40.15959-15964. DOI: 10.1073/pnas.0707912104.

Lévy, J., C. Bres, R. Geurts, B. Chalhoub, et al. (2004). "A Putative Ca2+ and Calmodulin-Dependent Protein Kinase for Bacterial and Fungal Symbioses". In: *Science* 303, pp. 1361–1363. DOI: DOI:10.1126/science.1093038.

Lhuissier, F. G. P., N. C. A. De Ruijter, B. J. Sieberer, J. J. Esseling, et al. (2001). "Time Course of Cell Biological Events Evoked in Legume Root Hairs by Rhizobium Nod Factors: State of the Art". In: *Annals of Botany* 87, pp. 289–302.

Li, B., C. Byrt, J. Qiu, U. Baumann, et al. (2016). "Identification of a Stelar-Localized Transport Protein That Facilitates Root-to-Shoot Transfer of Chloride

in Arabidopsis". In: *Plant Physiology* 170.1014-1029. DOI: 10.1104/pp.15.01163.

Li, D., E. Flores-Sandoval, U. Ahtesham, A. Coleman, et al. (2020). "Ethylene-independent functions of the ethylene precursor ACC in Marchantia polymorpha". In: *Nature Plant* 6, pp. 1335–1344. DOI: 10.1038/s41477-020-00784-y.

Li, J. F., J. E. Norville, M. Aach J .and McCormack, D. Zhang, et al. (2013). "Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9". In: *Nature Biotechnology* 31, pp. 688–691. DOI: 10.1038/nbt.2654.

Li, L., B. G. Kim, Y. H. Cheong, G. K. Pandey, et al. (2006). "A Ca2+ signaling pathway regulates a K+ channel for low-K response in Arabidopsis". In: *Proceedings of the National Academy of Science* 103, pp. 12625–12630.

Li, R., J. Zhang, J. Wei, H. Wang, et al. (2009). "Functions and mechanisms of the CBL–CIPK signaling system in plant response to abiotic stress". In: *Natural Science* 19, pp. 667–676. DOI: doi:10.1016/j.pnsc.2008.06.030.

Lin, H., Y. Yang, R. Quan, I. Mendoza, et al. (2009). "Phosphorylation of SOS3-LIKE CALCIUM BINDING PROTEIN8 by SOS2 protein kinase stabilizes their protein complex and regulates salt tolerance in Arabidopsis". In: *The Plant Cell* 21.5. DOI: 10.1105/tpc.109.066217.

Lindsay, M. P., E. S. Lagudah, and R. Hare R. A. Munns (2004). "A locus for sodium exclusion (Nax1), a trait for salt tolerance, mapped in durum wheat." In: *Functional Plant Biology* 31.1105-1114. DOI: doi.org/10.1071/FP04111.

Liu, H., H. M. Lyu, K. Zhu, Y. Van de Peer, et al. (2020). "The emergence and evolution of intron-poor and intronless genes in intron-rich plant gene families". In: *The Plant Journal* 105.4. DOI: 10.1111/tpj.15088.

Liu, K. H., Y. Niu, M. Konshi, Y. Wu, et al. (2017). "Discovery of nitrate-CPK-NLP signalling in central nutrient-growth networks". In: *Nature* 545.7654, pp. 311–316. URL: http://www.nature.com/articles/nature22077.pdf.

Liu, L. L., H. M. Ren, L. Q. Chen, Y. Wang, et al. (2013). "A Protein Kinase, Calcineurin B-Like Protein-Interacting Protein Kinase9, Interacts with Calcium Sensor Calcineurin B-Like Protein3 and Regulates Potassium Homeostasis under Low-Potassium Stress in Arabidopsis". In: *Plant Physiology* 161.1, pp. 266–277. ISSN: 0032-0889. DOI: 10.1104/pp.112.206896. eprint: http://www.plantphysiol.org/content/161/1/266.pdf. URL: http://www.plantphysiol.org/content/161/1/266.

Long, L., D. D. Guo, W. Gao, W. W. Yang, et al. (2018). "Optimization of CRISPR/-Cas9 genome editing in cotton by improved sgRNA expression." In: *Plant Methods* 14, p. 85. DOI: 10.1186/s13007-018-0353-0.

Lopez-Obando, M., R. de Villiers, B. Hoffmann, L. Ma, et al. (2018). "Physcomitrella patens MAX2 characterization suggests an ancient role for this F-box protein in photomorphogenesis rather than strigolactone signalling". In: *New Phytologist*, pp. 743–756. DOI: 10.1111/nph.15214.

Lu, Y., H. Hidaka, and J. Feldman (1996). "Characterization of a calcium/calmodulin-dependent protein kinase homolog from maize roots showing light-regulated gravitropism". In: *Planta* 199.1, pp. 18–24.

Luan, S., J. Kudla, M. Rodriguez-Concepcion, S. Yalovsky, et al. (2002). "Calmodulins and calcineurin B-like proteins: calcium sensors for specific signal response coupling in plants." In: *Plant Cell*. DOI: 10.1105/tpc.001115.

Lyzenga, W. J., V. Sullivan, H. Liu, and S. L. Stone (2017). "The Kinase Activity of Calcineurin B-like Interacting Protein Kinase 26 (CIPK26) Influences Its Own Stability and that of the ABA-regulated Ubiquitin Ligase, Keep on Going (KEG)". In: *Frontiers of Plant Science* 8, p. 502. DOI: 10.3389/fpls.2017.00502.

Ma, Y., L. Wang, J. Wang, Y. Zhong, et al. (2019). "Isolation and expression analysis of Salt Overly Sensitive gene family in grapevine (Vitisvinifera) in response to salt and PEG stress". In: *PLOS ONE* 14.e0212666. DOI: 10.1371/journal.pone.0212666.

Macnair, M. R., V. Bert, S. B. Huitson, P. Saumitou-Laprade, et al. (1999). "Zinc tolerance and hyperaccumulation are genetically independent characters". In: *The Royal Society* 266, pp. 2175–2179. DOI: 10.1098/rspb.1999.0905.

Maksaev, G. and E. S. Haswell (2012). "MscS-Like10 is a stretch-activated ion channel from Arabidopsis thaliana with a preference for anions". In: *Proceedings of National Academic Science* 109, pp. 19015–19020.

Mao, J., S. Manik, S. Shi, J. Chao, et al. (2016). "Mechanisms and Physiological Roles of the CBL-CIPK Networking System in Arabidopsis thaliana". In: *Genes* 7.9, p. 62.

Mao, Y., H. Zhang, N. Xu, B. Zhang, et al. (2013). "Application of the CRISPR-Cas system for efficient genome engineering in plants". In: *Molecular Plant* 6, pp. 2008–2011. DOI: 10.1093/mp/sst121..

Marks, R., J. Burton, and N. McLetchie (2016). "Sex differences and plasticity in dehydration tolerance: insight from a tropical liverwort". In: *Annals of Botany* 118, pp. 347–356. DOI: 10.1093/aob/mcw102.

Martinac, B. (2004). "Mechanosensitive ion channels: Molecules of Mechanotransduction". In: *Journal of Cell Science* 117, pp. 2449–2460.

McAinsh, M. R., C. Brownlee, and A. M. Hetherington (1990). "Abscisic acid-induced elevation of guard cell cytosolic Ca2+ precedes stomatal closure." In: *Nature* 343, pp. 186–188.

McCormack, M. and J. Braam (2003). "Calmodulins and related potential calcium sensors of Arabidopsis". In: *New Phytologist* 159.3, pp. 585–598. DOI: 10.1046/j.1469-8137.2003.00845.x.

Mehlmer, N., B. Wurzinger, S. Stael, D. Hofman-Rodrigues, et al. (2010). "The Ca2+-dependent protein kinase CPK3 is required for MAPK-independent salt-stress acclimation in Arabidopsis". In: *Plant Journal* 63, pp. 484–498. DOI: 10.1111/j.1365-313X.2010.04257.x.

Miller, J. B., A. Pratap, A. Miyahara, L. Zhou, et al. (2013). "Calcium/Calmodulin-Dependent Protein Kinase Is Negatively and Positively Regulated by Calcium,

Providing a Mechanism for Decoding Calcium Responses during Symbiosis Signaling". In: *The Plant Cell* 25.12. DOI: https://doi.org/10.1105/tpc.113.116921.

Miller, M. W. and J. Colaiace (1969). "The Induction of Sexual Reproductive Structures of Marchantia polymorpha Grown under Aseptic Culture Conditions". In: *The Bryologist* 72.1. DOI: 10.1639/0007-2745(1969)72[45:TIOSRS]2.0.CO;2.

Mithöfer, A. and C. Mazars (2002). "Aequorin-based measurements of intracellular Ca2+-signatures in plant cells". In: *Biological Procedures Online* 4.105-118. DOI: 10.1251/bpo40.

Monte, I., J. M. Franco-Zorrilla, G.-C. G., A. M. Zamarren, et al. (2019). "A Single JAZ Repressor Controls the Jasmonate Pathway in Marchantia polymorpha". In: *Molecular Plant* 12, pp. 185–198. DOI: 10.1016/j.molp.2018.12.017.

Morgan, A. J. and A. Galione (2014). "Teo-pore channels (TPCs): current controversies". In: *Bioessays* 36, pp. 173–183.

Munemasa, S., M. A. Hossain, Y. Nakamura, I. Mori, et al. (2010). "The Arabidopsis Calcium-Dependent Protein Kinase, CPK6, Functions as a Positive Regulator of Methyl Jasmonate Signaling in Guard Cells". In: *Plant Physiology* 155.1, pp. 553–561. DOI: 10.1104/pp.110.162750.

Munns, R. (2005). "Genes and Salt Tolerance: Bringing them together." In: *New Phytologist* 167.645-663. DOI: 10.1111/j.1469-8137.2005.01487.x.

Murata, Y., Z. M. Pei, I. C. Mori, and J. Schroeder (2001). "Abscisic Acid Activation of Plasma Membrane Ca2+ Channels in Guard Cells Requires Cytosolic NAD(P)H and Is Differentially Disrupted Upstream and Downstream of Reactive Oxygen Species Production in abi1-1 and abi2-1 Protein Phosphatase 2C Mutants." In: *The Plant Cell* 13, pp. 2513–2523.

Myers, C., S. M. Romanowsky, Y. D. Barron, S. Garg, et al. (2009). "Calcium-dependent protein kinases regulate polarized tip growth in pollen tubes". In: *Plant Journal* 59.4, pp. 528–539. DOI: doi:10.1111/j.1365-313X.2009.03894.x..

Nagata, T., S. Iizumi, K. Satoh, H. Ooka, et al. (2004). "Comparative Analysis of Plant and Animal Calcium Signal Transduction Element Using Plant Full-Length cDNA Data". In: *Molecular Biology and Evolution* 21.10.

Nakayama, N., H. Kawasaki, and R. H. Kretsinger (2000). "Evolution of EF-hand proteins." In: *Topics in Biological Inorganic Chemistry* 3.29-58. DOI: https://doi.org/10.1007/978-3-642-58306-3_2.

Negae, M., A. Nozawa, N. Koizumi, H. Sano, et al. (2003). "The crystal structure of the novel calcium-binding protein AtCBL2 from Arabidopsis thaliana". In: *Journal of Biological Chemistry* 278.43, pp. 42240–42246. DOI: DOI:10.1074/jbc.M303630200.

Nekrasov, V., B. Staskawicz, D. Weigel, J. D. Jones, et al. (2013). "Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease". In: *Nature Biotechnology* 31, pp. 691–693. DOI: 10.1038/nbt.2655.

Nichloas, M., G. Marais, V. Hykelova, B. Janousek, et al. (2005). "A Gradual Process of Recombination Restriction in the Evolutionary History of the Sex Chromosomes in Dioecious Plants". In: *PLOS Biology* 3.1, e4. DOI: https://doi.org/10.1371/journal.pbio.0030004.

Nieves-Cordones, M., F. Caballero, V. Martinez, and F. Rubio (2012). "Disruption of the Arabidopsis thaliana Inward-Rectifier K+ Channel AKT1 Improves Plant Responses to Water Stress". In: *Plant and Cell Physiology* 53.423-432. DOI: 10.1093/pcp/pcr194.

Nimigean, C. M., T. Shane, and C. Miller (2004). "A cyclic nucleotide modulated prokaryotic K+ channel". In: *Journal of Genetic Physiology* 124, pp. 203–210.

Nishiyama, R., S. Mizuno H. Okada, T. Yamaguchi, M. Takenaka, et al. (1999). "Two mRNA species encoding calcium-dependent protein kinases are differentially expressed in sexual organs of Marchantia polymorpha through alternative splicing." In: *Plant and Cell Physiology* 40.2, pp. 205–212.

O'Hanlon, M. E. (1926). "Germination of the spores and early stages in thedevelopment of the gametophyte ofMarchantia polymorpha." In: *Botany Gazette* 82, pp. 215–222.

Oda, K., K. T. Yamato, E. Ohta, Y. Nakamura, et al. (1992). "Gene organization deduced from the complete sequence of liverwort Marchantia polymorpha mitochondrial DNA. A primitive form of plant mitochondrial genome." In: *Journal of Molecular Biology* 223, pp. 1–7.

Oh, S. I., J. Park, S. Yoon, Y. Kim, et al. (2008). "The Arabidopsis Calcium Sensor Calcineurin B-Like 3 Inhibits the 5'-Methylthioadenosine Nucleosidase in a Calcium-Dependent Manner". In: *Plant Physiology* 148.4. DOI: 10.1104/pp.108.130419.

Ohta, M., Y. Guo, U. Halfter, and J. K. Zhu (2003). "A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2". In: *Proceedings of the National Academy of Sciences of the United States of America* 100.11771-11776.

Ohyama, K., H. Fukuzawa, T. Kohchi, H. Shirai, et al. (1986). "Chloroplast Gene Organization Deduced from Complete Sequence of Liverwort Marchantia polymorpha Chloroplast DNA." In: *Nature* 322, pp. 572–574.

Okazaki, Y., M. Ishigami, and N. Iwasaki (2002). "Temporal Relationship between Cytosolic Free Ca2+ and Membrane Potential during Hypotonic Turgor Regulation in a Brackish Water Charophyte Lamprothamnium succinctum." In: *Plant & Cell Physiology* 43, pp. 1027–1035.

Pandey, G. K., J. J. Grant, Y. H. Cheong, B. G. Kim, et al. (2008). "Calcineurin-B-Like Protein CBL9 Interacts with Target Kinase CIPK3 in the Regulation of ABA Response in Seed Germination". In: *Molecular Plant* 1.2, pp. 238–248. ISSN: 1674-2052. DOI: https://doi.org/10.1093/mp/ssn003. URL: http://www.sciencedirect.com/science/article/pii/S1674205214604329.

Pei, Z. M., Y. Murata, G. Benning, S. Thomine, et al. (2000). "Calcium channels activated by hydrogen peroxide mediate abscisic acid signalling in guard cells". In: *Nature* 406.731-734.

Peiter, E., M. Fischer, K. Sidaway, S. K. Roberts, et al. (2005a). "The Saccharomyces cerevisiae ca2+ channel Cch1pMid1p is essentisal for tolerance to cold stress and iron toxicity". In: *FEBS Letters* 579, pp. 5697–5703.

Peiter, E., F. J. M. Maathuis, L. N. Mills, H. Knight, et al. (2005b). "The vacuolar Ca2+-activated channel TPC1 regulates germination and stomatal movement". In: *Nature* 434, pp. 404–408.

Pleith, A. and A. Trewavas (2002). "Reorientation of seedlings in the earths gravitational field induces cytosolic calcium transients". In: *Plant Physiology* 129, pp. 786–796.

Popescu, S. C., G. V. Popescu, S. Bachan, Z. Zhang, et al. (2007). "Differential binding of calmodulin-related proteins to their targets revealed through high-density Arabidopsis protein microarrays". In: *PNAS* 104.11, pp. 4730–4735. DOI: https://doi.org/10.1073/pnas.0611615104.

Pretty, J. (2007). "Agricultural Sustainability: Concepts, Principles and Evidence". In: *Philos Trans R Soc Lond B Biol Sci* 1491.447-465. DOI: 10.1098/rstb.2007.2163.

Puttick, M. N., J. L. Morris, T. A. Williams, C. J. Cox, et al. (2018). "The Interrelationships of Land Plants and the Nature of the Ancestral Embryophyte." In: *Current Biology* 5, pp. 733–745. DOI: 10.1016/j.cub.2018.01.063..

Qiu, Q. S., Y. Guo, M. A. Dietrich, K. S. Schmaker, et al. (2002). "Regulation of SOS1, a plasma membrane Na+/H+ exchanger in Arabidopsis thaliana, by SOS2 and SOS3." In: *Proceedings from the National Academy of Science* 99.2, pp. 8436–8441.

Quan, R., H. Lin, I. Mendoza, Y. Zhang, et al. (2007). "SCABP8/CBL10, a Putative Calcium Sensor, Interacts with the Protein Kinase SOS2 to Protect Arabidopsis

Shoots from Salt Stress". In: *The Plant Cell* 19, pp. 1415–1431. DOI: https://doi.org/10.1105/tpc.106.042291.

Quintero, F. J., J. Martinez-Atienza, I. Villalta, X. Jiang, et al. (2011). "Activation of the plasma membrane Na/H antiporter Salt-Overly-Sensitive 1 (SOS1) by phosphorylation of an auto-inhibitory C-terminal domain." In: *Proceedings from the National Academy of Science* 108.6, pp. 2611–2616.

Quintero, F. J., M. Ohta, H. Shi, J. K. Zhu, et al. (2002). "Reconstitution in yeast of the Arabidopsis SOS signaling pathway for Na+ homeostasis." In: *Proceedings from the National Academy of Science* 99.1, pp. 9061–9066.

Ragel, P., R. Rodenas, E. Garcia-Martin, Z. Andres, et al. (2015). "The CBL-Interacting Protein Kinase CIPK23 Regulates HAK5-Mediated High-Affinity K+ Uptake in Arabidopsis Roots". In: *Plant Physiology* 169.4, pp. 2863–2873.

Ramachandiran, S., D. Takezawa, W. Wand, and B. W. Poovaiah (1997). "Functional Domains of Plant Chimeric Calcium/Calmodulin-Dependent Protein Kinase: Regulation by Autoinhibitory and Visinin-Like Domains". In: *The Journal of Biochemistry* 121.5, pp. 984–990. DOI: https://doi.org/10.1093/oxfordjournals.jbchem.a021684.

Ranf, S., P. Wunnenberg, J. Lee, D. Becker, et al. (2008). "Loss of vacuolar cation channel, AtTPC1, does not impair calcium signals induced by abiotic and biotic stresses". In: *Plant Journal* 53, pp. 287–299.

Reddy, V. S. and A. S. N. Reddy (2004). "Proteomics of calcium-signaling components in plants". In: *Phytochemistry* 65.12, pp. 1745–1776. DOI: 10.1016/j.phytochem.2004.04.033.

Ren, X., G. Qi, H. Feng, S. Zhao, et al. (2013). "Calcineurin B-like protein CBL10 directly interacts with AKT1 and modulates K+ homeostasis in Arabidopsis". In: *The Plant Journal* 74.2, pp. 258–266.

Réty, S., D. Osterloh, J. P. Arié, S. Tabaries, et al. (2000). "Structural basis of the Ca(2+)-dependent association between S100C (S100A11) and its target, the N-

terminal part of annexin I". In: *Structure* 15.8. DOI: 10.1016/s0969-2126(00)00093-9..

Rodenas, R. and G. Vert (2021). "Regulation of Root Nutrient Transporters by CIPK23: 'One Kinase to Rule Them All'". In: *Plant and Cell Physiology* 62.4. DOI: 10.1093/pcp/pcaa156.

Rolly, N. K., Q. M. Imran, I. J. Lee, and B. W. Yun (2020). "Salinity Stress-Mediated Suppression of Expression of Salt Overly Sensitive Signaling Pathway Genes Suggests Negative Regulation by AtbZIP62 Transcription Factor in Arabidopsis thaliana". In: *International Journal of Molecular Sciences* 21.1726. DOI: 10.3390/ijms21051726.

Ronzier, E., C. Corratgé-Faillie, F. Sanchez, K. Prado, et al. (2014). "CPK13, a noncanonical Ca2+-dependent protein kinase, specifically inhibits KAT2 and KAT1 shaker K+ channels and reduces stomatal opening." In: *Plant Physiology* 166, pp. 314–326. DOI: 10.1104/pp.114.240226.

Rost, B. (1999). "Twilight zone of protein sequence alignments". In: *Protein Engineering, Design, Selection* 12.2, pp. 85–94. DOI: https://doi.org/10.1093/protein/12.2.85.

Rubinigg M. Wenisch, J., J. T. Elzenga, and I. Stulen (2004). "NaCl salinity affects lateral root development in Plantago maritima". In: *Functional Plant Biology* 31.775-780. DOI: 10.1071/FP03222.

Rudd, J. J., F. C. H. Franklin, J. M. Lord, and V. E. Flanklin-Tong (1996). "Increased Phosphorylation of a 26-kD Pollen Protein Is Induced by the Self-Incompatibility Response in Papaver rhoeas". In: *Plant Cell* 8.4, pp. 713–724. DOI: 10.1105/tpc.8.4.713.

Ruszala, E. M., D. J. Beerling, P. J. Franks, C. Chater, et al. (2011). "Land plants acquired active stomatal control early in their evolutionary history". In: *Current Biology* 21, pp. 1030–1035. DOI: 10.1016/j.cub.2011.04.044.

Saint-Marcoux, D., H. Proust, L. Dolan, and J. A. Langdale (2015). "Identification of Reference Genes for Real-Time Quantitative PCR Experiments in the Liverwort

Marchantia polymorpha". In: *PLOS ONE* 10, e0118678. DOI: 10.1371/journal.pone.0118678.

Sánchez-Barrena, M. J., A. Chaves-Sanjuan, N. Raddatz, I. Mendoza, et al. (2020). "Recognition and Activation of the Plant AKT1 Potassium Channel by the Kinase CIPK23". In: *Plant Physiology* 182, pp. 2143–2153. DOI: doi.org/10.1104/pp.19.01084.

Sánchez-Barrena, M. J., H. Fujii, I. Angulo, M. Martínez-Ripoll, et al. (2007). "The structure of the C-terminal domain of the protein kinase AtSOS2 bound to the calcium sensor AtSOS3". In: *Molecular Cell* 26.3. DOI: 10.1016/j.molcel.2007.04.013..

Sánchez-Barrena, M. J., M. Martínez-Ripoll, J. Zhu, and A. Albert (2005). "The structure of the Arabidopsis thaliana SOS3: molecular mechanism of sensing calcium for salt stress response." In: *Journal of Molecular Biology* 345.5. DOI: 10.1016/j.jmb.2004.11.025.

Sánchez-Barrena, M., M. Martínez-Ripoll, and A. Albert (2013). "Structural Biology of a Major Signaling Network that Regulates Plant Abiotic Stress: The CBL-CIPK Mediated Pathway". In: *International Journal of Molecular Science* 14.3, pp. 5734–5749. DOI: doi:10.3390/ijms14035734.

Sanyal, A. K., P. Kanwar, A. K. Yadav, C. Sharma, et al. (2017a). "Arabidopsis CBL interacting protein kinase 3 interacts with ABR1, an APETALA2 domain transcription factor, to regulate ABA responses". In: *Plant Science* 254, pp. 48–59. DOI: 10.1016/j.plantsci.2016.11.004.

Sanyal, S. K., P. Kanwar, H. Samtani, K. Kaur, et al. (2017b). "Alternative Splicing of CIPK3 Results in Distinct Target Selection to Propagate ABA Signaling in Arabidopsis". In: *Frontiers in Plant Science* 8, p. 1924. ISSN: 1664-462X. DOI: 10.3389/fpls.2017.01924. URL: https://www.frontiersin.org/article/10.3389/fpls.2017.01924.

Schmid, M. W., A. Giraldo-Fonseca, M. Rovekamp, D. Smetanin, et al. (2018). "Extensive epigenetic reprogramming during the life cycle of Marchantia poly-

morpha". In: *Genome Biology* 19.9. DOI: https://doi.org/10.1186/s13059-017-1383-z.

Schmöckel, S. M., A. F. Garcia, B. Berger, M. Tester, et al. (2015). "Different NaCl-Induced Calcium Signatures in the Arabidopsis thaliana Ecotypes Col-0 and C24". In: *PLOS ONE* 2.e0117564. DOI: 10.1371/journal.pone.0117564.

Seki, M., M. Narusaka, J. Ishida, T. Nanjo, et al. (2002). "Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray." In: *The Plant Journal* 31.279-292. DOI: 10.1046/j.1365-313X.2002.01359.x.

Shao, Q., Q. Gao, D. Lhamo, H. Zhang, et al. (2020). "Two glutamate- and pH-regulated Ca2+ channels are required for systemic wound signaling in Arabidopsis". In: *Science Signalling* 13, eaba1453. DOI: 10.1126/scisignal.aba1453.

Sharma, N., P. L. Bhalla, and M. B. Singh (2013). "Transcriptome-wide profiling and expression analysis of transcription factor families in a liverwort, Marchantia polymorpha". In: *BMC Genomics* 23.14, p. 915.

Sharma, N., C. H. Jung, P. L. Bhalla, and M. B. Singh (2014). "RNA Sequencing Analysis of the Gametophyte Transcriptome from the Liverwort, Marchantia polymorpha". In: *PLOS ONE* 9, e97497. DOI: 10.1371/journal.pone.0097497.

Shi, H., L. Xiong, B. Stevenson, T. Lu, et al. (2002). "The Arabidopsis salt overly sensitive 4 mutants uncover a critical role for vitamin B6 in plant salt tolerance." In: *Plant Cell* 14.3, pp. 575–588.

Shi, J., K. N. Kim, V. Ritz O. Albrecht, R. Gupta, et al. (1999). "Novel Protein Kinases Associated with Calcineurin B–like Calcium Sensors in Arabidopsis". In: *The Plant Cell* 11.12. DOI: 10.1105/tpc.11.12.2393.

Shimamura, M. (2016). "Marchantia polymorpha: Taxonomy, Phylogeny and Morphology of a Model System". In: *Plant Cell Physiology* 57.2, pp. 230–256.

Sillitoe, I., A. L. Cuff, B. H. Dessailly, N. L. Dawson, et al. (2013). "New functional families (FunFams) in CATH to improve the mapping of conserved functional

sites to 3D structures." In: *Nucleic Acids Res.* 41 (Database issue).D490-8. DOI: 10.1093/nar/gks1211.

Silva, A. C. da, J. Kendrick-Jones, and F. C. Reinach (1995). "Determinants of ion specificity on EF-hands sites. Conversion of the Ca2+/Mg2+ site of smooth muscle myosin regulatory light chain into a Ca(2+)-specific site." In: *Journal of Biological Chemistry* 270.6773-8 (12). DOI: 10.1074/jbc.270.12.6773.

Sirichandra, C., D. Gu, H. C. Hu, M. Davanture, et al. (2009). "Phosphorylation of the Arabidopsis AtrbohF NADPH oxidase by OST1 protein kinase." In: *FEBS Lett.* 583.2982-2986. DOI: 10.1016/j.febslet.2009.08.033..

Stephan, A. B. and J. Schroeder (2014). "Plant salt stress status is transmitted systemically via progating calcium waves". In: *Proceedings from the National Academy of Science* 111, pp. 6126–6127.

Su, D., L. Yand, X. Shi, X. Ma, et al. (2021). "Large-Scale Phylogenomic Analyses Reveal the Monophyly of Bryophytes and Neoproterozoic Origin of Land Plants". In: *Nature Plants* 38, pp. 3332–3344. DOI: 10.1093/molbev/msab106.

Sugano, S. S., R. Nishihama, M. Shirakawa, J. Takagi, et al. (2018). "Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in Marchantia polymorpha". In: *bioRxiV*. DOI: https://doi.org/10.1101/277350.

Sugano, S. S., M. Shirakawa, J. Takagi, Y. Matsuda, et al. (2014). "CRISPR/Cas9-mediated targeted mutagenesis in the liverwort Marchantia polymorpha L". In: *Plant Cell Physiol* 55.3, pp. 475–81. ISSN: 1471-9053 (Electronic) 0032-0781 (Linking). DOI: 10.1093/pcp/pcu014. URL: http://www.ncbi.nlm.nih.gov/pubmed/24443494.

Sun, X., Z. Hu, R. Chen, Q. Jiang, et al. (2015). "Targeted mutagenesis in soybean using the CRISPR-Cas9 system". In: *Scientific Reports* 5, p. 10342. DOI: 10.1038/srep10342.

Swainsbury, D. J. K., L. Zhou, G. E. D. Oldroyd, and S. Bornemann (2012). "Calcium Ion Binding Properties of Medicago truncatula Calcium/Calmodulin-Dependent Protein Kinase". In: *Biochemistry* 51.35, pp. 6895–6907.

Szebenyi, D. M. E. and K. Moffat (1986). "The refined structure of vitamin D-dependent calcium-binding protein from bovine intestine. Molecular detail, ion binding, and implicaions for the structure of other calcium-binding proteins." In: *Journal of Biological Chemistry* 261.8761-8777.

Takahashi, H. and Y. Kodama (2020). "CRUNC: a cryopreservation method for unencapsulated gemmae of Marchantia polymorpha". In: *PeerJ*, e10174. DOI: 10.7717/peerj.10174..

Takenaka, M., S. Yamaoka, T. Hanajiri, Y. Shimizu-Ueda, et al. (2000). "Direct transformation and plant regeneration of the haploid liverwort Marchantia polymorpha L." In: *Transgenic. Re.* 9.3, pp. 179–185.

Talke, I. N., D. Blaudez, F. J. M. Maathuis, and D. Sanders (2003). "CNGCs: Prime targets of plant cyclic nucleotide signalling?" In: *Trends in Plant Science* 8, pp. 286–293.

Tanaka, D., K. Ishizaki, T. Kohchi, and K. T. Yamato (2015). "Cryopreservation of Gemmae from the Liverwort Marchantia polymorpha L." In: *Plant and Cell Physiology* 57.2, pp. 300–306. DOI: https://doi.org/10.1093/pcp/pcv173.

Tanaka, H., R. Suzuki, N. Okabe, T. Suzuki, et al. (2018). "Salinity stress-responsive transcription factors in the liverwort Marchantia polymorpha". In: *Plant Biotechnology* 35.281-284. DOI: 10.5511/plantbiotechnology.18.0501a.

Tang, R., H. Liu, Y. Yang, L. Yang, et al. (2012). "Tonoplast calcium sensors CBL2 and CBL3 control plant growth and ion homeostasis through regulating V-ATPase activity in Arabidopsi". In: *Cell Research* 22.12, pp. 1650–1665.

Thor, K., S. Jiang, E. Michard, J. George, et al. (2020). "The calcium-permeable channel OSCA1.3 regulates plant stomatal immunity". In: *Nature* 585, pp. 569–573. DOI: 10.1038/s41586-020-2702-1.

Tomescu, A. M. F., B. Bomfleur, A. C. Bippus, and A. Savoretti (2018). "Chapter 16 - Why Are Bryophytes So Rare in the Fossil Record? A Spotlight on Taphonomy and Fossil Preservation". In: *Transformative Paleobotany*, pp. 375–416. DOI: 10.1016/B978-0-12-813012-4.00016-4.

Toyota, M., H. Furuichi, H. Tatsumi, and M. Sokabe (2008). "Cytoplasmic calcium increases in response to changes in the gravity vector in hypocotyls and petioles of Arabidopsis seedlings". In: *Plant Physiology* 146, pp. 505–514.

Trewavas, A. and M. R. Knight (1994). "Mechanical signalling: Calcium and Plant form". In: *Plant Molecular Biology* 26, pp. 1329–1341.

Tsuboyama, S. and Y. Kodama (2014). "AgarTrap: a simplified Agrobacterium-mediated transformation method for sporelings of the liverwort Marchantia polymorpha L." In: *Plant Cell Physiology* 55, pp. 229–236. DOI: 10.1093/pcp/pct168..

Tsuboyama, S., S. Nonaka, H. Ezura, and Y. Kodama (2018). "Improved G-AgarTrap: A highly efficient transformation method for intact gemmalings of the liverwort Marchantia polymorpha". In: *Scientific Reports* 8, p. 10800. DOI: 10.1038/s41598-018-28947-0.

Tsuboyama-Tanaka, S. and Y. Kodama (2015). "AgarTrap-mediated genetic transformation using intact gemmae/gemmalings of the liverwort Marchantia polymorpha L". In: *J Plant Res* 128.2, pp. 337–44. ISSN: 1618-0860 (Electronic) 0918-9440 (Linking). DOI: 10.1007/s10265-014-0695-2. URL: http://www.ncbi.nlm.nih.gov/pubmed/25663453.

Tutar, Y. (2012). "Psuedogenes". In: *International Journal of Genomics*. DOI: 10.1155/2012/424526.

Valmonte, G. R., K. Arthur, C. M. Higgins, and R. M. MacDiarmid (2014). "Calcium-dependent protein kinases in plants: evolution, expression and function." In: *Plant Cell Physiology* 55, pp. 551–569. DOI: 10.1093/pcp/pct200.

Veley, K. M., S. Marshburn, C. E. Clure, and E. S. Haswell (2012). "Mechanosensitive channels protect plastids from hypoosmotic stress during normal plant growth," in: *Current Biology* 22, pp. 408–413.

Vera-Estrella, R., M. F. Gomez-Mendez, J. C. Amezcua-Romero, B. J. Barkla, et al. (2017). "Cadmium and zinc activate adaptive mechanisms in Nicotiana tabacum similar to those observed in metal tolerant plants". In: *Planta* 246, pp. 433–451. DOI: 10.1007/s00425-017-2700-1.

Verret, F., G. Wheeler, A. R. Taylor, G. Farnham, et al. (2010). "Calcium Channels in Photosynthetic Eukaryotes: Implications for evolution of Calcium-based signalling". In: *New Phytologist* 187, pp. 79–89.

Vincent, T. R., M. Avramova, J. Canham, P. Higgins, et al. (2017). "Interplay of Plasma Membrane and Vacuolar Ion Channels, Together with BAK1, Elicits Rapid Cytosolic Calcium Elevations in Arabidopsis during Aphid Feeding." In: *Plant Cell* 29.6, pp. 1460–1479. DOI: 10.1105/tpc.17.00136.

Waadt, R., L. K. Schmidt, M. Lohse, K. Hashimoto, et al. (2008). "Multicolor bimolecular fluorescence complementation reveals simultaneous formation of alternative CBL/CIPK complexes in planta". In: *Plant Journal* 56, pp. 505–516.

Wagner, S., S. Behera, S. D. Bortoli, D. C. Logan, et al. (2015). "The EF-Hand Ca2+ Binding Protein MICU Choreographs Mitochondrial Ca2+ Dynamics in Arabidopsis". In: *Plant Cell* 27.11, pp. 3190–3212. DOI: 10.1105/tpc.15.00509.

Wagner, S., S. D. Bortoli, M. Schwarzlander, and I. Szabo (2016). "Regulation of mitochondrial calcium in plants versus animals". In: *Journal of Experimental Biology* 67.13, pp. 3809–3829. DOI: https://doi.org/10.1093/jxb/erw100.

Wang, L., S. Liang, and Y. Lu (2001). "Characterization, physical location, and expression of the genes encoding calcium/calmodulin-dependent protein kinases in maize". In: *Planta* 2001.213, p. 4.

Wang, P. H., C. E. Lee, Y. S. Lin, M. H. Lee, et al. (2019). "The Glutamate Receptor-Like Protein GLR3.7 Interacts With 14-3-3ω and Participates in Salt Stress

Response in Arabidopsis thaliana". In: *Frontiers in Plant Science*. DOI: 10.3389/fpls.2019.01169.

Wang, X., X. Zhang, X. P. Dong, M. Samie, et al. (2012). "TCP proteins and phosphoinositide-activated sodium-selective ion channels in endosomes and lysosomes". In: *Cell* 151, pp. 372–383.

Wang, Y., S. Liang, Q. G. Xie, and Y. T. Lu (2004). "Characterization of a calmodulin-regulated Ca2+-dependent-protein-kinase-related protein kinase, AtCRK1, from Arabidopsis". In: *Biochemical Journal* 282.Pt1, pp. 73–81. DOI: 10.1042/BJ20031907.

Wang, Y. F., S. Munemasa, N. Nishimura, H. M. Ren, et al. (2013). "Identification of cyclic GMP-activated nonselective Ca2+-permeable cation channels and associated CNGC5 and CNGC6 genes in Arabidopsis guard cells". In: *Plant Physiology* 163, pp. 578–590.

Ward, J. M. and J. Schroeder (1994). "Calcium activated K+ channels and calcium induced calcium release by slow vacuolar ion channels in guards cell vacuoles implicated in the control of stomatal closure". In: *Plant Cell* 6, pp. 669–683.

Watillon, B., R. Kettmann, P. Boxus, and A. Burny (1992). "Cloning and characterization of an apple (Malus domestica (L.) Borkh) cDNA encoding a calmodulin-binding protein domain similar to the calmodulin-binding region of type II mammalian Ca2+/calmodulin-dependent protein kinase". In: *Plant Science* 81.2, pp. 227–235. DOI: DOI:10.1016/0168-9452(92)90046-O.

Watillon, B., R. Kettmann, P. Boxus, and A. Burny (1993). "A calcium/calmodulin-binding serine/threonine protein kinase homologous to the mammalian type I1 calcium/calmodulin-dependent bprotein kinase is expressed in plant cells". In: *Plant Physiology* 101, pp. 1381–1384.

Watillon, B., R. Kettmann, P. Boxus, and A. Burny (1995). "Structure of a calmodulin-binding protein kinase gene from apple". In: *Plant Physiology* 108.2, pp. 847–848.

Webb, A. A. R., M. R. McAinsh, J. E. Taylor, and A. M. Hetherington (1996). "Calcium ions as intracellular second messengers in higher plants". In: *Advances*

in Botanical Research, Vol 22 22, pp. 45–96. ISSN: 0065-2296. URL: %3CGo%20to%20ISI%3E://WOS:A1996BF36W00002.

Weckwerth, P., B. Ehlert, and T. Romeis (2014). "ZmCPK1, a calcium independent kinase member of the Zea mays CDPK gene family, functions as a negative regulator in cold stress signalling". In: *Plant, Cell and Environment* 38, pp. 544–558. DOI: 10.1111/pce.12414.

Weinl, S. and J. Kudla (2009). "The CBL-CIPK Ca2+-decoding signaling network: function and perspectives". In: *New Phytologist* 184.3, pp. 517–528.

Wen, F., F. Ye, Z. Xiao, L. Liao, et al. (2020). "Genome-wide survey and expression analysis of calcium-dependent protein kinase (CDPK) in grass Brachypodium distachyon". In: *BMC Genomics* 21, p. 53. DOI: 10.1186/s12864-020-6475-6.

Wernimont, A., J. D. Artz, P. Finerty, Y. Lin, et al. (2010). "Structures of apicomplexan calcium-dependent protein kinases reveal mechanism of activation by calcium". In: *Nature Structural and Molecular biology* 17.5, pp. 596–601. DOI: doi:10.1038/nsmb.1795.

Whalley, H. J. and M. R. Knight (2013). "Calcium signatures are decoded by plants to give specific gene responses." In: *New Phytologist* 197.3, pp. 690–693.

Whalley, H. J., A. W. Sargeant, J. F. Steele, T. Lacoere, et al. (2011). "Transcriptomic analysis reveals calcium regulation of specific promoter motifs in Arabidopsis." In: *Plant Cell* 23.11, pp. 4079–4095.

Wheeler, G. and C. Brownlee (2008). "Ca2+ signalling in plants and green algae-changing channels". In: *Trends in Plant Science* 13.506-514.

Wilson, C. H., E. S. Ali, N. Scrimgeour, A. M. Martin, et al. (2015). "Steatosis inhibits liver cell store-operated Ca2+ entry and reduces ER Ca2+ through a protein kinase C-dependent mechanism". In: *Biochemical Journal* 466, pp. 379–390. ISSN: 0264-6021. DOI: 10.1042/bj20140881. URL: %3CGo%20to%20ISI%3E://WOS:000351686800017.

Wu, T. Y., H. Goh, C. B. Azodi, S. Krishnamoorthi, et al. (2021). "Evolutionarily conserved hierarchical gene regulatory networks for plant salt stress response". In: *Nature Plants* 7.787-799. DOI: 10.1038/s41477-021-00929-7.

Xiao, F., X. Li, J. He, J. Zhao, et al. (2021). "Protein kinase PpCIPK1 modulates plant salt tolerance in Physcomitrella patens." In: *Plant Molecular Biology* 105, pp. 685–696. DOI: 10.1007/s11103-021-01120-4..

Xie, K. and Y. Yang (2013). "RNA-guided genome editing in plants using a CRISPR-Cas system". In: *Molecular Plant* 6, pp. 1975–1983. DOI: 10.1093/mp/sst119.

Xing, T., X. J. Wang, K. Malik, and B. L. Miki (2001). "Ectopic Expression of an Arabidopsis Calmodulin-Like Domain Protein Kinase-Enhanced NADPH Oxidase Activity and Oxidative Burst in Tomato Protoplasts". In: *APS* 14.10, pp. 1261–1264. DOI: https://doi.org/10.1094/MPMI.2001.14.10.1261.

Xu, J., Y. S. Tian, R. H. Pen, A. S. Xiong, et al. (2010a). "Cyanobacteria MT gene SmtA enhance zinc tolerance in Arabidopsis". In: *Molecular Biology Reports* 37, p. 1105. DOI: 10.1007/s11033-009-9867-x.

Xu, J., Y. S. Tian, R. H. Peng, A. S. Xiong, et al. (2010b). "AtCPK6, a functionally redundant and positive regulator involved in salt/drought stress tolerance in Arabidopsis". In: *Planta* 231.1251-1260. DOI: 10.1007/s00425-010-1216-8.

Yamato, K. T., K. Ishizaki, M. Fujisawa, S. Okada, et al. (2007). "Gene Organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system". In: *Proceedings of the National Academy of Science* 104, pp. 6472–6477.

Yang, J., P. T. Ellinor, W. A. Sather, J. F. Zhang, et al. (1993). "Molecular determinants of Ca2+ selectivity and ion permeation in L-type Ca2+ channels". In: *Nature* 366, pp. 158–161.

Yang, Q., Z. Z. Chen, X. F. Zhou, H. B. Yin, et al. (2009). "Overexpression of SOS (Salt Overly Sensitive) genes increases salt tolerance in transgenic Arabidopsis". In: *Molecular Plant* 2, pp. 22–31. DOI: 10.1093/mp/ssn058.

Yang, T. and B. W. Poovaiah (2003). "Calcium/calmodulin-mediated signal network in plants". In: *Trends in Plant Science* 8.10, pp. 505–512. DOI: DOI:10.1016/j.tplants.2003.09.004.

Yin, X., Y. Xia, Q. Xie, Y. Cao, et al. (Dec. 2019). "The protein kinase complex CBL10–CIPK8–SOS1 functions in Arabidopsis to regulate salt tolerance". In: *Journal of Experimental Botany* 71.6, pp. 1801–1814. ISSN: 0022-0957. DOI: 10.1093/jxb/erz549. eprint: https://academic.oup.com/jxb/article-pdf/71/6/1801/33282637/erz549.pdf. URL: https://doi.org/10.1093/jxb/erz549.

Yu, L., J. Nie, C. Cao, Y. Jin, et al. (2010). "Phosphatidic acid mediates salt stress response by regulation of MPK6 in Arabidopsis thaliana". In: *New Phytologist* 188.3, pp. 762–773.

Yu, X., M. Li, G. Gao, H. Feng, et al. (2006). "Abscisic acid stimulates a calcium-dependent protein kinase in grape berry". In: *Plant Physiology* 140, pp. 558–579.

Yuan, F., H. Yang, Y. Xue, D. Kong, et al. (2014). "OSCA1 mediates osmotic-stress-evoked Ca2+ increases vital for osmosensing in Arabidopsis." In: *Nature* 514.7522, pp. 367–371. DOI: 10.1038/nature13593.

Zhang, H., B. Yang, W. Z. Liu, H. Li, et al. (2014). "Identification and characterization of CBL and CIPK gene families in canola (Brassica napus L.)" In: *BMC Plant Biology* 14, p. 8. DOI: 10.1186/1471-2229-14-8.

Zhang, H., Y. Zhang, C. Deng, S. Deng, et al. (2018). "The Arabidopsis Ca2+-Dependent Protein Kinase CPK12 Is Involved in Plant Response to Salt Stress". In: *International Journal of Molecular Sciences* 19.4062. DOI: 10.3390/ijms19124062.

Zhang, J., X. X. Fu, R. Q. Li, X. Zhao, et al. (2020). "The hornwort genome and early land plant evolution." In: *Nature Plants* 6, pp. 107–118. DOI: 10.1038/s41477-019-0588-4.

Zhang, L. and Y. T. Lu (2003). "Calmodulin binding protein kinases in plants." In: *Trends in Plant Science* 8.3, pp. 123–127. DOI: 10.1016/S1360-1385(03)00013-X.

Zhang Y. Wang, Y., J. L. Taylor, S. Jiang Z. Zhang, F. Mei, et al. (2015). "Aequorin-based luminescence imaging reveals differential calcium signalling responses to salt and reactive oxygen species in rice roots". In: *Journal of Experimental Botany* 66.2535-2545. DOI: 10.1093/jxb/erv043.

Zhao, R., H. Sun, N. Zhao, X. Jing, et al. (2015). "The Arabidopsis Ca2+-dependent protein kinase CPK27 is required for plant response to salt-stress". In: *Gene* 563, pp. 203–214. DOI: 10.1016/j.gene.2015.03.024.

Zhou, L., W. Lan, Y. Jiang, W. Fang, et al. (2014). "A calcium-dependent protein kinase interacts with and activates a calcium channel to regulate pollen tube growth". In: *Molecular Plant* 163, pp. 369–376.

Zhu, J. K., J. Liu, and L. Xiong (1998). "Genetic Analysis of Salt tolerance in Arabidopsis: Evidence for a Critical Role of Potassium Nutrition". In: *The Plant Cell* 10, pp. 1181–1191. DOI: 10.1105/tpc.10.7.1181.

Zhu, K., F. Chen, L. J., X. Chen, et al. (2016). "Evolution of an intron-poor cluster of the CIPK gene family and expression in response to drought stress in soybean." In: *Scientific Reports* 6.28225. DOI: https://doi.org/10.1038/srep28225.

Zhu, S. Y., X. C. Yu, X. J. Wang, R. Zhao, et al. (2007). "Two calcium-dependent protein kinases, CPK4 and CPK11, regulate abscisic acid signal transduction in Arabidopsis". In: *Plant Cell* 19.3019-3036. DOI: 10.1105/tpc.107.050666.

Zhu, X., C. Dunand, W. Snedden, and J. P. Galaud (2015). "CaM and CML emergence in the Green Lineage". In: *Trends in Plant Science* 20.8, pp. 483–489. DOI: https://doi.org/10.1016/j.tplants.2015.05.010.

Zhu, X., A. Taylor, S. Zhang, D. Zhang, et al. (2014). "Measuring spatial and temporal Ca2+ signals in Arabidopsis plants". In: *J Vis Exp* 2, e51945. DOI: 10.3791/51945.

Zou, J. H., F. J. Wei, J. J. Wang C. Wu, D. Ratnasekera, et al. (2010). "Arabidopsis Calcium-Dependent Protein Kinase CPK10 Functions in Abscisic Acid- and Ca2+-Mediated Stomatal Regulation in Response to Drought Stress". In: *Plant Physiology* 154.1232-1243. DOI: 10.1104/pp.110.157545.